GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 20, 2002, 11:35:32 ; Search time 21 Seconds
(without alignments)
2485.761 Million cell updates/sec

Title: US-09-759-207-2
Perfect score: 2842
Sequence: 1 MLLRSKPALPPPLMLLLLGP.....LPAFSYSFFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	897.5	31.6	480	2	JC7506	heparanase protein
2	416	14.6	521	2	T45608	hypothetical prote
3	169.5	6.0	190	2	T01953	hypothetical prote
4	112.5	4.0	356	2	F64383	hypothetical prote
5	111.5	3.9	575	2	T12094	beta-fructofuranos
6	111	3.9	670	2	T10666	hypothetical prote
7	111	3.9	688	2	S32961	hypothetical prote
8	111	3.9	2298	2	T49648	hypothetical prote
9	109.5	3.9	879	2	E91031	probable outer mem
10	108.5	3.8	411	2	S74760	hypothetical prote
11	107.5	3.8	500	2	D87541	beta-xylosidase [i
12	106	3.7	879	2	F85875	probable fimbrial
13	105	3.7	670	2	T38446	microtubule-associ
14	104.5	3.7	788	1	S00652	phosphoribosylamin
15	104	3.7	432	2	F70411	adenylosuccinate s
16	104	3.7	2013	2	AI1489	probable peptidogl
17	103.5	3.6	587	2	S36231	beta-fructofuranos
18	103.5	3.6	676	2	AF1153	transcription anti
19	103.5	3.6	687	2	F85188	retrotransposon li
20	103	3.6	796	2	D97065	transketolase [imp
21	101	3.6	594	2	A82913	hypothetical prote
22	101	3.6	644	2	A97268	methionyl-tRNA syn
23	100.5	3.5	805	2	C86525	DNA gyrase subunit
24	100.5	3.5	805	2	H72098	DNA gyrase, chain
25	100.5	3.5	989	2	AE2140	toxin secretion AB
26	99.5	3.5	510	2	H69893	conserved hypothet
27	99.5	3.5	837	1	A31842	endo-1,4-beta-xyla
28	99	3.5	897	2	G02529	dynein heavy chain
29	99	3.5	4644	1	A38905	dynein heavy chain
30	98.5	3.5	596	2	T04506	hypothetical prote
31	98.5	3.5	629	2	C64180	hypothetical prote
32	98.5	3.5	654	2	T14202	NADH2 dehydrogenas
33	98.5	3.5	699	2	F95146	DNA topoisomerase
34	98.5	3.5	701	2	D98014	DNA topoisomerase
35	98.5	3.5	746	2	T46821	siderophore recept
36	98.5	3.5	746	2	A95420	RhtA Rhizobactin r
37	98.5	3.5	1012	2	JC5925	membrane klotho pr
38	98	3.4	465	2	T19113	hypothetical prote
39	98	3.4	716	1	C60008	RNA-directed RNA p
40	98	3.4	760	2	T34414	hypothetical prote
41	98	3.4	817	2	H75035	probable membrane
42	97.5	3.4	454	2	T20829	probable serine ca
43	97.5	3.4	511	2	S61166	probable membrane
44	97.5	3.4	604	2	E75119	hypothetical prote
45	97.5	3.4	804	2	G71546	probable DNA gyras

ALIGNMENTS

RESULT 1
JC7506
heparanase protein 2a - human
C;Species: Homo sapiens (man)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 01-Dec-2000
C;Accession: JC7506
R;McKenzie, E.; Tyson, K.; Stamps, A.; Smith, P.; Turner, P.; Barry, R.; Hircock, Biochem. Biophys. Res. Commun. 276, 1170-1177, 2000
A;Title: Cloning and expression profiling of Hpa2, a novel mammalian heparanase fa
A;Reference number: JC7506
A;Accession: JC7506
A;Molecule type: mRNA
A;Residues: 1-480 <MCK>
A;Cross-references: GB:AF282885
C;Comment: This protein, a intracellular membrane-bound enzyme, has biological and therapies.
C;Genetics:
A;Gene: hpa2a
A;Map position: 10q23-10q24
C;Keywords: heparin binding; membrane bound

Query Match 31.6%; Score 897.5; DB 2; Length 480;
Best Local Similarity 36.0%; Pred. No. 2.6e-59;
Matches 202; Conservative 74; Mismatches 146; Indels 139; Gaps 9;

```
Qy      20 PLGPLSPGAL-----------------PRPA------QAQDVVDLDFFTQEPLHLVSPS 55
           |   |:||||                 | |       : : :: ||  |: |:  |: :
Db      18 PPACLAPGALYLALLLHLSLSSQAGDRRPLPVDRAAGLKEKTLILLDVSTKNPVRTVNEN 77

Qy      56 FLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIF----DPKKEST 111
           |||: :| ::  |  :|  | | :| |||||||||:||||| :|||| |    :| |
Db      78 FLSLQLDPSIIHD-GWLDFLSSKRLVTLARGLSPAFLRFGGKRTDFLQFQNLRNPAKSR- 135

Qy     112 FEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVD 171
                         |   ||                   :| | ::
Db     136 ------------------GGPGPD------------------YYLKNYE---------- 148

Qy     172 VLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLK 231
                                                              :|||::
Db     149 ----------------------------------------------------DEPNNYRT 156

Qy     232 KADIFINGSQLGEDYIQLHKLLRK-STFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVI 290
              :||||||:|||||  ||:    :  | ||||::|:||:    :|  |:|  |  :
Db     157 MHGRAVNGSQLGKDYIQLKSLLQPIRIYSRASLYGPNIGRPRKNVIALLDGFMKVAGSTV 216

Qy     291 DSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGG 350
           |:||| | |::||     |||   :||     ::|: :|| :  ||||:||    :   |
Db     217 DAVTWQHCYIDGRVVKVMDFLKTRLLDTLSDQIRKIQKVVNTYTPGKKIWLEGVVTTSAG 276

Qy     351 GAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLF 410
           |   |||::||||:||: ||: |  ||:||:|  ||  |  ||||:||:||||||||||:
Db     277 GTNNLSDSYAAGFLWLNTLGMLANQGIDVVIRHSFFDHGYNHLVDQNFNPLPDYWLSLLY 336
```

```
Qy    411 KKLVGTKVLMASVQGSKRR---------KLRVYLHCTNTDNPRYKEGDLTLYAINLHNVT 461
Db    337 KRLIGPKVLAVHVAGLQRKPRPGRVIRDKLRIYAHCTNHHNHNYVRGSITLFIINLHRSR 396

Qy    462 KYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSS 521
Db    397 KKIKLAGTLRDKLVHQYLLQPYGQEGLKSKSVQLNGQPLVMVDDGTLPELKPRPLRAGRT 456

Qy    522 LGLPAFSYSFFVIRNAKVAAC 542
Db    457 LVIPPVTMGFFVVKNVNALAC 477
```

RESULT 2
T45608
hypothetical protein F13G24.30 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T45608
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23009
A;Accession: T45608
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-521 <BEV>
A;Cross-references: EMBL:AL133421
A;Experimental source: cultivar Columbia; BAC clone F13G24
C;Genetics:
A;Map position: 5
A;Introns: 53/3; 66/1; 127/2; 177/1; 256/1; 319/2; 361/2; 394/3
A;Note: F13G24.30

  Query Match 14.6%; Score 416; DB 2; Length 521;
  Best Local Similarity 29.2%; Pred. No. 3.3e-23;
  Matches 154; Conservative 68; Mismatches 184; Indels 122; Gaps 24;

```
Qy     75 LGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIPP 134
Db     55 LTRPLLTKAIKAFKPLRIRIGGSLQDQVIYDVGNLKT--------------PCR------ 94

Qy    135 DVEEKLRLEWPYQEQLLLREHYQKKFKNS---TYSRSSV-----DVLYTFANCSGLDLIF 186
Db     95 ----------PFQKM------------NSGLFGFSKGCLHMKRWDELNSFLTATGAVVTF 132

Qy    187 GLNALLRTADLQ-------WNSSNAQLLLDYCSSKGYNI-SWELGNEPNSFLKKADIFIN 238
Db    133 GLNALRGRHKLRGKAWGGAWDHINTQDFLNYTVSKGYVIDSWEFGNELSG--SGVGASVS 190

Qy    239 GSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQP-----RRKTAKMLKSFLKAGGEVIDSV 293
Db    191 AELYGKDLIVLKDVINK-VYKNSWLHKPILVAPGGFYEQQWYTKLLEI---SGPSVVDVV 246

Qy    294 TWHHYYLNGRT--ATREDFLNPDVLDIFISSVQKVF----QVVESTRPGKKVWLGETSSA 347
Db    247 THHIYNLGSGNDPALVKKIMDPS----YLSQVSKTFKDVNQTIQEHGPWASPWVGESGGA 302

Qy    348 YGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDE-NFDPLPDYWL 406
Db    303 YNSGGRHVSDTFIDSFWYLDQLGMSARHNTKVYCRQTLVG-GFYGLLEKGTFVPNPDYYS 361

Qy    407 SLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYL-- 464
Db    362 ALLWHRLMGKGVLAVQTDGPP--QLRVYAHCSK------GRAGVTLLLINLSNQSDFTVS 413

Qy    465 ---------------------RLPYPFS---NKQVDKYLLRP---LGPHG--LLSKSVQL 495
Db    414 VSNGINVVLNAESRKKKSLLDTLKRPFSWIGSKASDGYLNREEYHLTPENGVLRSKTMVL 473

Qy    496 NGLTLKMVDDQTLPPLMEKPLRP-GSSLGLPAFSYSFFVIRNAKVAAC 542
Db    474 NGKSLKPTATGDIPSL-EPVLRSVNSPLNVLPLSMSFIVLPNFDASAC 520
```

RESULT 3
T01953
hypothetical protein T2L5.6 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jan-2000
C;Accession: T01953
R;Geisel, C.; Smith, A.; Le, T.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of A. thaliana T2L5.
A;Reference number: Z14470
A;Accession: T01953
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-190 <GEI>
A;Cross-references: EMBL:AF096371; NID:g3695386; PID:g3695392
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 36/2; 69/3
A;Note: T2L5.6
C;Superfamily: Arabidopsis thaliana hypothetical protein T2L5.6

  Query Match 6.0%; Score 169.5; DB 2; Length 190;
  Best Local Similarity 27.8%; Pred. No. 2.1e-05;
  Matches 54; Conservative 34; Mismatches 57; Indels 49; Gaps 9;

```
Qy    382 RQVFFGAGNYHLVD-ENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNT 440
Db     12 RQSLIG-GNYGLLNTTNFTPNPDYYSALIWRQLMGRKALFTTFSGTK--KIRSYTHCA-- 66

Qy    441 DNPRYKEGDLTLYAINLHNV----------------TKYLRLPYPFSNKQVDKYLLRPL 483
Db     67 ---RQSKG-ITVLLMNLDNTTTVVAKVELNNSFSLRHTKHMK-----SYKRASSQLFG-- 115

Qy    484 GPHGLL---------------SKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFS 528
Db    116 GPNGVIQREEYHLTAKDGNLHSQTMLLNGNALQVNSMGDLPPIEPIHINSTEPITIAPYS 175

Qy    529 YSFFVIRNAKVAAC 542
Db    176 IVFVHMRNVVVPAC 189
```

RESULT 4
F64383
hypothetical protein MJ0670 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: F64383
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.;
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Gloc
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; V
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jar
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: F64383
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-356 <BUL>
A;Cross-references: GB:U67514; GB:L77117; NID:g2826304; PIDN:AAB98664.1; PID:g15
C;Genetics:
A;Map position: REV596956-595886
A;Start codon: GTG

  Query Match 4.0%; Score 112.5; DB 2; Length 356;
  Best Local Similarity 21.2%; Pred. No. 1;
  Matches 85; Conservative 48; Mismatches 143; Indels 125; Gaps 18;

```
Qy    126 ICKY-----GSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVD--------- 171
```

```
Db     14 JRKYMKIYNGKNEKDIKERLIKE-------LKEEHVLVETEDGTYTLKAEDEEEMMHSKV 66

Qy    172 -----VLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEP 226
Db     67 GALKEAIYKFAKPS-----------KITDL----SNPR-VLDLCSGMGYNAIAALHYNK 109

Qy    227 NS----------------FL------------KKADIFINGSQLGEDYIQLHKLLRKSTF 258
Db    110 NAEIDMVEICEEVLFLTLFLDIPYKEHEIIKDKVREYFLN--KIGIEY--------KSDY 159

Qy    259 KNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTAT--REDFLNPDVL 316
Db    160 DNINLY---VGDARKFIIKSDKKY---------NVVFHDAFSPKRDPTLYTYDFL----- 202

Qy    317 DIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMG 376
Db    203 -------KEIYKRMEDN--GVLI-------SYSSAIPFRSALVDCGFVISEKESVGRKRG 246

Qy    377 IEVVMRQVFFGAGNYHLVDENFD-------PLPDYWLSLLFKKLVGTKVLMASVQGSKRR 429
Db    247 ITLAYKNPNFKPNRINEVDERVIALSVIALPYRDETLSLTKDKIIEDREERREKLKEKLI 306

Qy    430 KLRVYLHCTNTDNPRYKEGDLTLYA--INLHNVTKYLRLPY 468
Db    307 KIGKYLSTKQIKKGNIPEEILKIQKEDLNSSEIIKKMRLKF 347


RESULT 5
T12094
beta-fructofuranosidase (EC 3.2.1.26) - fava bean
C;Species: Vicia faba (fava bean)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T12094
R;Weber, H.; Borisjuk, L.; Heim, U.; Buchner, P.; Wobus, U.
Plant Cell 7, 1835-1846, 1995
A;Title: Seed coat-associated invertases of Fava bean control both unloading and storage
A;Reference number: Z17416; MUID:96093423; PMID:8535137
A;Accession: T12094
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-575 <WEB>
A;Cross-references: EMBL:Z35162; NID:g861154; PIDN:CAA84526.1; PID:g861155
A;Experimental source: cv. Fribo, seed coat
C;Genetics:
A;Gene: CWINV1
C;Superfamily: beta-fructofuranosidase
C;Keywords: cell wall; glycoprotein; glycosidase; hydrolase

  Query Match             3.9%;  Score 111.5;  DB 2;  Length 575;
  Best Local Similarity   21.4%;  Pred. No. 2.5;
  Matches   72;  Conservative   48;  Mismatches  107;  Indels  109;  Gaps   19;

Qy     46 QEPLHLVS-------PSFLSVTIDANLATDPRFLILLGSPKLRTLARGLS---------P 89
Db    228 KHPIHSAKRTGMWECPDFYPVSLEGKNGLD--LSMMMGNNVKHVLKNSLDITRYEYYTIG 285

Qy     90 AYLR--------------FGGTKTDF-------LIFDPKKESTFEERSYW----QSQVNQ 124
Db    286 TYLQNQDKYIPDKTSEDGWGGLRYDYGNFYASKSFFDPTK----NRRIIWGWANESDTKE 341

Qy    125 DICKYG-----SIPPDV-----EEKLRLEWPYQEQLLLR----EHYQKKFKNSTYSRSSV 170
Db    342 DDVKKGWAGIQAIPRTVWLDSSRRQLR-QWPVEELNRLRGKQVEMKNRKLKKGGY----L 396

Qy    171 DVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFL 230
Db    397 EVKGITASQADVEVTFSFSSLDKAEAFDPNWENAE---DLCAQKGSKVRGGVG--PFGLL 451

Qy    231 KKADIFINGSQLGEDYIQL----------HKLL----RKSTFKNAKLYGP--------DV 268
Db    452 TLA------SKKLEEYTSVFFRVFKAANKHKILMCSDAKSSSLNRELYKPSFAGFVNVDL 505

Qy    269 GQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRT 304
```

```
Db    506 GNNKKLSLRSL----------IDHSVVESFGVGGKT 531


RESULT 6
T10666
hypothetical protein F6E21.40 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
C;Accession: T10666
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.;
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16533
A;Accession: T10666
A;Molecule type: DNA
A;Residues: 1-670 <BEV>
A;Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.40
A;Experimental source: cultivar Columbia; BAC clone F6E21
C;Genetics:
A;Gene: ATSP:F6E21.40
A;Map position: 4
A;Introns: 47/3; 87/1; 123/3; 203/3; 230/2; 255/3; 284/3; 305/1; 335/3; 347/3; 370/
C;Superfamily: Schizosaccharomyces pombe negative regulator of mitosis skb1

  Query Match             3.9%;  Score 111;  DB 2;  Length 670;
  Best Local Similarity   22.4%;  Pred. No. 3.4;
  Matches  123;  Conservative   77;  Mismatches  194;  Indels  156;  Gaps   33;

Qy     51 LVSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAY--LRFGGTKTDFLIFDPK- 107
Db     47 LVDPSYRPSLVEGN-GVDTQVLPVCGSDLV------LSPSQWSSHVVGKISSWIDLDSED 99

Qy    108 ------KESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFK 161
Db    100 EVLRMDSETTLKQEIAWATHLSLQMCE-----PD---------------LTRPHYLAGGL 139

Qy    162 NSTYSRSSV----DVLY--------TFANCS--GLDLIFGLNALLRTADLQWNSSNAQLL 207
Db    140 RVSCCRSSFISDETFLYKITFNQALTFCGSSLFCLNVISALKLWLRVPLVK---SEGDSM 196

Qy    208 LDYCSSKGYNISWELGN------EPNSFLKKA-DIFIN-------GSQLGEDYIQLHKLL 253
Db    197 DD--TSEGLNDSWELWNSFRLLCEHDSKLSVALDVLSTLPSETSLGRWMGES-VRAAILS 253

Qy    254 RKSTFKNAKLYGPDVGQP--RRKTAKMLKSFL-KAGGEVIDSVTWHHYYLNGRTATREDF 310
Db    254 TDAFLTNAR------GYPCLSKRHQKLIAGFFDHAAQVVICGKPVHNLQKPLDSSSEGTE 307

Qy    311 LNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLG 370
Db    308 KNP--LRIYLDYVAYLFQKMESLSEQERIELGYRDFLQAPLQPLMDNLEAQTYETFE--- 362

Qy    371 LSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM---------- 420
Db    363 ---RDSVKYIQYQ---RAVEKALVDR----VPDEKASEL------TTVLMVVGAGRGPLV 406

Qy    421 -ASVQGSKR--RKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTK---------YLRLPY 468
Db    407 RASLQAAEETDRKLKVY---AVEKNPN--------AVVTLHNLVKMEGWEDVVTIISCDM 455

Qy    469 PFSN--KQVDKYLLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLM---EKPLRPGSSLG 523
Db    456 RFWNAPEQADILVSELLGSFG----------------DNELSPECLDGAQRFLKP-DGIS 498

Qy    524 LPAFSYSFFV 533
Db    499 IPS-SYTSFI 507


RESULT 7
S32961
hypothetical protein YBR259w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YBR1727
```

```
C;Species: Saccharomyces cerevisiae
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-Apr-2002
C;Accession: S32961; S46140
R;Doignon, F.; Biteau, N.; Crouzet, M.; Aigle, M.
Yeast 9, 189-199, 1993
A;Title: The complete sequence of a 19,482 bp segment located on the right arm of chromo
A;Reference number: S29348; MUID:93220397; PMID:8465606
A;Accession: S32961
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-688 <DOI>
A;Cross-references: EMBL:X70529; NID:g1907246; PIDN:CAA49923.1; PID:g296558
R;Aigle, M.; Baclet, M.C.; Barthe, C.; Biteau, N.; Crouzet, M.; Doignon, F.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45940
A;Accession: S46140
A;Molecule type: DNA
A;Residues: 1-688 <AIG>
A;Cross-references: EMBL:Z36128; NID:g536684; PIDN:CAA85222.1; PID:g536685; MIPS:YBR259w
C;Genetics:
A;Cross-references: SGD:S0000463
A;Map position: 2R
C;Superfamily: Saccharomyces cerevisiae hypothetical protein YBR259w

  Query Match             3.9%;  Score 111;  DB 2;  Length 688;
  Best Local Similarity   22.5%;  Pred. No. 3.5;
  Matches   67;  Conservative   45;  Mismatches   94;  Indels   92;  Gaps   16;

Qy     126 ICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSS----------VDVLYT 175
           : :| |   | ::| :|:: |: :: |:|   | ::|    :||           ::|
Db     164 MAEYSSWKWDSDDKRQLQFMYEFRMKLKECLVKFYENFDLQKSSDPLKELIIPWEKIVYV 223

Qy     176 FANCSGLDLIFGLNALLRTADLQWNSSN--------AQLLLD-----YCSSKGY------ 216
           |||  :|   |    :  |:| | | |        | | |:    : : : : |
Db     224 -ANC--IDAFTGEQVRIDGAELIWTSKNLVFSSISSAVLRLNDLQNMFSAFRPYGEEALV 280

Qy     217 --------NISWELGNEPNSFLKKA---DIF--INGSQLG--EDYIQLHKLLRK------- 255
               :: |:  ::  | ::     |:|   |  |:    | |   :||||
Db     281 QDFAHIRSLKWDSNDKVESLIRALIFNDMFPYFNKEQVDTKADGIFFLRLLRKNFKEHIN 340

Qy     256 --------------STFKN--AKLYGPDVGQPRRKTAKMLKSFLKAGGEV------IDSV 293
                        | |||  : |      | |||:  |  | |  | ::      ||
Db     341 DVKDFHIQVIKYLNSQFKNNYSTLMTSSKTQDRRKSHNMPSSILDDGNKIGMHVSPIDE- 399

Qy     294 TWHHYYLNG----------RTATREDFLNPDVLDIFISSVQKVFQVVESTR---PGKK 338
           : |:  |           :  | |    ||   || |   |:: :: '  |   | |:
Db     400 -YSHFIDNDEPLWRDKVYPKIYTNEQTPTPDASAIFDS--HKIYAIISLLRYYLPEKR 454


RESULT 8
T49648
hypothetical protein B8B20.20 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49648
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49648
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2298 <SCH>
A;Cross-references: EMBL:AL355933; GSPDB:GN00116; NCSP:B8B20.20
A;Experimental source: BAC clone B8B20; strain OR74A
C;Genetics:
A;Gene: NCSP:B8B20.20
A;Map position: 6
A;Introns: 426/3

  Query Match             3.9%;  Score 111;  DB 2;  Length 2298;
  Best Local Similarity   19.3%;  Pred. No. 22;
  Matches  114;  Conservative   79;  Mismatches  190;  Indels  208;  Gaps   28;

Qy      78 PKLRTLARGLSPA---------YLRFGGTKTDFLIFDPKKESTFEERSYWQSQV-NQDIC 127
           |::  : | ::|:          :|       :: : :  ::|     |:: | || :
Db    1447 PRVDLIERLITPSNAHKEACLINIRAWNQLARLVVSNGEGSASFRPFITWRNNVFNQILD 1506

Qy     128 KYGSIPPDVEEKLR---------LEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFAN 178
           :| |   |:|:: |         ::  ::|:|: :               :::|:|:| |
Db    1507 QYMSAESDIEQQFRALSAENMRSIDAAWREELITKN----------KATALDILHTSAR 1555

Qy     179 CSGLDLIFGLNAL-----LRTADLQ-------------WNSSNAQLLLDYCSSKGYNI 218
           | ||::      |      |    ||               |   | |  |      : : :
Db    1556 AS-LDVLKQAKTLEAAIYTLNVTQLQKMCTTLHFGSPGFDWGILNVAL-----DTHAHFL 1609

Qy     219 SW-ELGNEPNSFLKKADIFINGSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAK 277
           | |  :|      ::   |:  || :  : | : | |  |  |      |   |
Db    1610 GWIETSSEEQYSSNESSADIDPRQLEDAILLLQEKLTKEFFWMA----------RELLAL 1659

Qy     278 MLKSFLKAGGEVID-SVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPG 336
           ||:    | :    : |     |  : | |           ||   ::| ||:   :||
Db    1660 PLKAITTFGKQTEQVACTEKTVTLAAKLAAR----------FIQ--ERVTQVLPYFQPG 1706

Qy     337 K---------------KVWLGETSSAYGGGAPLLSDTFAAG------------FMWLDK 368
           |                : ||          ||   |                :|:
Db    1707 KYGLFPDMPKNMSGPERRWL----------PLFIATLVNKNVFDFKDIETNILSLWVQS 1755

Qy     369 LGLSAR-MGIEVVMRQVFFGAGNYHL--VDENFDPLPDYWLSL-LFKKLVG--TKVL--- 419
           :    | :| |  : :|     |   |  | :    ||| : | || : :    | |
Db    1756 IIKPMRFLGYETYLAEVLQQRGLPFLAEADVSAGMTPDYNIHLDLFSRAIHYMRKALRGG 1815

Qy     420 ----------MASVQGSK-----RRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYL 464
                      ||  ||      ||:   : |
Db    1816 ATTPAPGVTSSASTAGSSSAQSIRRQREEFSH---------------------------- 1847

Qy     465 RLPYPFSNKQVDKYLLRPLG-----------------PHGLLSKSVQLNGLTLKMVD-- 504
           |    :| : | :|||                    |||:|  :: :|: : :||
Db    1848 TLQLAMTNIKKDLFLLRSLALADPTASSTEEHRDYMAFTHGLIS-LIKSHGVGIVVVDSF 1906

Qy     505 -----DQTLPPLMEKPLRPG---------SSLGLPAFSYSF-FVIRNAKVA 540
                |   ||| :  |              |   :|| |  | :: | |||
Db    1907 FLTPSDSYSPPLQDPQLHTAGIMAYGVRLSEKDVPAASQLFWYLFNNFKVA 1957


RESULT 9
E91031
probable outer membrane protein ECs3221 [imported] - Escherichia coli (strain O15
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: E91031
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Ha
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 a
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E91031
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-879 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36644.1; PID:g13362691; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs3221

  Query Match             3.9%;  Score 109.5;  DB 2;  Length 879;
  Best Local Similarity   20.1%;  Pred. No. 6.7;
  Matches  130;  Conservative   68;  Mismatches  208;  Indels  241;  Gaps   33;

Qy      52 VSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKEST 111
           :| |:::   :  :  | ||| | |  |:  | |  |  |:  |  |  :  :   |:  |
Db      20 MSGSYVNAWAENEIQFDSRFLELKGDTKI-DLKRFSSQGYVEPG--KYNLQVQLNKQPLT 76

Qy     112 FEERSYWQSQVNQDICKYGSIPPDVEEKL--------RLEWPYQEQLLLREHYQKKFKNS 163
           |   || :  |    | : |::  :          |:| :  : |       :
Db      77 EEYDIYWYASENDASKTYACLTPELVAQFGLKEDVAKNLQWIHDGKCLKPGQLE------ 130
```

```
Qy    164 TYSRSSVDVLYTFANCSGLDLIFGL-NALLRTADLQWNSSN------AQLLLDYC----- 211
Db    131 -----GIDIK---ADLSQSALVISLPQAYLEYTDINWDPPSRWDDGISGLIADYSITAQT 182

Qy    212 ----------------SSKGYNI-SWEL-GNEPNSFL---KKADIFINGSQ--------- 241
Db    183 RHEENGGDDSNEISGNGTVGVNLGAWRLRADWQTDYLHSKSNDDDVINGDDTQKNWEWSR 242

Qy    242 ----------------LGEDYIQLHKLLRKSTF-----------------KNAKLYGPDV 268
Db    243 YYAWRALPSLKAKLGLGEDY------LNSDIFDGFNYVGGSISTDDQMLPPNLRGYAPDI 296

Qy    269 GQPRRKTAKMLKSFLKAGGEVI---------------DSVTWHHYYLNGRTATREDFLN 312
Db    297 SGVAHTTAKVTVSQL---GRVIYETQVPAGPFRIQDLGDSV-------SGTLHIRIEEQN 346

Qy    313 PDVLDIFISSVQKVFQVVESTRPGK---KVWL----------------GETS------- 345
Db    347 GQVQEYDINTASMPF----LTRPGQVRYKLMMGRPQEWGHHVEGGFFSGGEASWGIANGW 402

Qy    346 SAYGGGAPLLSD---------------TFAA-------GFMWLDKL---------GLSAR 374
Db    403 SLYGGA---LADEHYQSAALGVGRDLSVFGAVAFDITHSHTRLDKETAYGKGSLDGNSFR 459

Qy    375 MGI-----EVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGT---KVLMASVQGS 426
Db    460 LSYSKDFDELNSRVTFAG---YRFSEENFMTMSEY-LDASDSEMVRTGNDKEMYTATYNQ 515

Qy    427 KRRKLRVYLHCTNTDNPRYKEGDLTLYAI------NLHNVTK----------------- 462
Db    516 NFRDAGVSVYLNYTRHTYWDRDEQTNYNVMLSHYFNLGSIRNMSISMTGYRYEYDNQADK 575

Qy    463 --YLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTLKMVDDQT 507
Db    576 GVYISLSMPWGDSSTISY----NGNYGSGSDSSQVG--YFSRVDDAT 616
```

RESULT 10
S74760
hypothetical protein slr1617 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: S74760
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74760
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <KAN>
A;Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAA16911.1; PID:d101764
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

  Query Match 3.8%; Score 108.5; DB 2; Length 411;
  Best Local Similarity 21.4%; Pred. No. 2.5;
  Matches 66; Conservative 55; Mismatches 94; Indels 93; Gaps 17;

```
Qy    159 KFKNSTYSRSSVDVLYTFANCSGLDLIF---GLNALLRTADLQWNSSNAQL--------- 206
Db    151 EFRLISPTREQIDI---FAGSTKLDLLASEENIDCIVHLANPRVYTSNVAMGQTLTMLRN 207

Qy    207 LLDYCSSKG----YNISWEL---------GNEPNSFLKK---------ADIFINGSQLGE 244
Db    208 VIDVCLAKDIPLIYPSSWEIYSGYAGTIHADESTPALPRGPYGETKYLAEILI------- 260

Qy    245 DYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRT 304
Db    261 DHCRRTRGLRCAILRSSPVYGSMSDKP-----KFIFNFFKKASQGQKIVT--HHYING-- 311

Qy    305 ATREDFLNPDV----LDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
Db    312 -------NPKLDLLHIDDLISSIVATL---------KSRFIGNLNI----GTGQLSSTLK 351

Qy    361 AGFMWLDKLGLSA-----RMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVG 415
Db    352 IAEMIRDELGSSSMIQQIEVNTEVASIAMNYGRAN-HVLD---------WEPVIFFE-QG 400

Qy    416 TKVLMASV 423
Db    401 LKSLLHQI 408
```

RESULT 11
D87541
beta-xylosidase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: D87541
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelbe B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H., n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87541
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-500 <STO>
A;Cross-references: GB:AE005673; NID:g13423886; PIDN:AAK24328.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC2357

  Query Match 3.8%; Score 107.5; DB 2; Length 500;
  Best Local Similarity 25.7%; Pred. No. 4;
  Matches 56; Conservative 35; Mismatches 80; Indels 47; Gaps 15;

```
Qy    165 YSRSSVDVLYTFANCSGLDLIFGLN---ALLRTAD---LQW--NSSNAQL-----LLD-- 209
Db     81 YDWTKIDQLYDALLAKGIKPFIELGFTPEAMKTSDQTIFYWKGNTSHPKLGPWRDLIDAF 140

Qy    210 -YCSSKGYNI----SW--ELGNEPN--SFLKKADIFINGSQLGEDYIQLHKLLRKSTFKN 260
Db    141 VHHLRARYGVEEVRTWFFEVWNEPNLDGFWEKAD--------QAAYFELYDV----TARA 188

Qy    261 AKLYGPD--VGQPRRKTAKMLKSFL---KAGGEVIDSVTWHHYYLNG----RTATREDFL 311
Db    189 IKAIDPSLRVGGPATAGAAWVPEFLAHVKKSGSAVDFVTTHTYGVDGGFLDEKGVQDTKL 248

Qy    312 NPDVLDIFISSVQKVFQVVE-STRPGKKVWLGETSSAY 348
Db    249 SPSP-DAVVGDVRRVREQIEASAFPGLPLYFTEWSTSY 285
```

RESULT 12
F85875
probable fimbrial usher Z3600 [imported] - Escherichia coli (strain O157:H7, substr
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85875
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; B iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apc Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-879 <STO>
A;Cross-references: GB:AE005174; NID:g12516702; PIDN:AAG57466.1; GSPDB:GN00145; UW(
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3600

```
  Query Match             3.7%;  Score 106;  DB 2;  Length 879;
  Best Local Similarity   20.5%;  Pred. No. 12;
  Matches  129;  Conservative   64;  Mismatches  201;  Indels  234;  Gaps   34;

Qy      52 VSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKEST 111
           :| |:::   :  :  | |||| | |  |:  | |  |  |:  |  | :  :   |:  |
Db      20 MSGSYVNAWAENEIQFDSRFLELKGDTKI-DLKRFSSQGYVEPG--KYNLQVQLNKQPLT 76

Qy     112 FEERSYWQSQVNQDICKYGSIPPDVEEKL--------RLEWPYQEQLLLREHYQKKFKNS 163
            |   || :  |     |  : |::  :          |:| :  : |      :
Db      77 EEYDIYWYASENDASKTYACLTPELVAQFGLKEDVAKNLQWIHDGKCLKPGQLE------ 130

Qy     164 TYSRSSVDVLYTFANCSGLDLIFGL-NALLRTADLQWNSSN------AQLLLDYC----- 211
                 :|:    |: |   |:  |  | |   |: |:  :      : |: ||
Db     131 -----GIDIK---ADLSQSALVPSLPQAYLEYTDINWDPPSRWDDGISGLIADYSITAQT 182

Qy     212 ----------------SSKGYNI-SWEL-GNEPNSFL---KKADIFINGSQ--------- 241
                           : | |  :| |  :    :|      |  |||
Db     183 RHEENGGDDSNEISGNGTVGVNXGAWRLRADWQTDYLHSKSNDDDVINGDDTQKNWEWSR 242

Qy     242 ---------------LGEDYIQLHKLLRKSTF-----------------KNAKLYGPDV 268
                           |||||      |    |                 | : | ||:
Db     243 YYAWRALPSLKAKLGLGEDY------LNSDIFDGFNYVGGSISTDDQMLPPNLRGYAPDI 296

Qy     269 GQPRRKTAKMLKSFLKAGGEVI----------------DSVTWHHYYLNGRTATREDFLN 312
                 |||:  | |   | ||                |||       :|    | :  |
Db     297 SGVAHTTAKVTVSQL---GRVIYETQVPAGPFRIQDLGDSV-------SGTLHIRIEEQN 346

Qy     313 PDVLDIFISSVQKVFQVVESTRPGK---KVWL----------------GETS------- 345
              | :  |::    |     ||||:   |: :                || |
Db     347 GQVQEYDINTASMPF----LTRPGQVRYKLMMGRPQEWGHHVEGGFFSGGEASWGIANGW 402

Qy     346 SAYGGGAPLLSD---------------TFAA-------GFMWLDKL---------GLSAR 374
           | |||    |:|                | |           |||           | | |
Db     403 SLYGGA---LADEHYQSAALGVGRDLSVFGAVAFDITHSHTRLDKETAYGKGSLDGNSFR 459

Qy     375 MGI-----EVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGT---KVLMASVQGS 426
           :       |:  |  | |   |   :|||  : :| |     ::| |   | :  :
Db     460 LSYSKDFDELNSRVTFAG---YRFSEENFMTMSEY-LDASDSEMVRTGNDKEMYTATYNQ 515

Qy     427 KRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRL------------PYPFSNK 473
           |   | ::   | :  :   : | | : |   : |  |                | : |
Db     516 NFRDAGVSVYLNYTRHTYWDRDEQTNYNVML---SHYFNLGSIRNMSISMTGYRYEYDN- 571

Qy     474 QVDK----YLLRPLGPHGLLLSKSVQLNG 497
           | ||     |  | |     | ::  ||
Db     572 QADKGVYISLXMPWGD----SSTISYNG 595


RESULT 13
T38446
microtubule-associated protein ssm4 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T38446; T00012
R;McDougall, R.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1997
A;Reference number: Z21793
A;Accession: T38446
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-670 <MCD>
A;Cross-references: EMBL:AL009227; PIDN:CAA15832.1; GSPDB:GN00066; SPDB:SPAC27D7.13c
A;Experimental source: strain 972h-; cosmid c27D7
R;Yamashita, A.; Watanabe, Y.; Yamamoto, M.
Genes to Cells 2, 155-166, 1997
A;Title: Microtubule-associated coiled-coil protein Ssm4 is involved in the meiotic deve
A;Reference number: Z14042; MUID:97311255; PMID:9167972
A;Accession: T00012
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-670 <YAM>
A;Cross-references: EMBL:AB000269; NID:g3341860; PIDN:BAA31857.1; PID:g3341861
C;Genetics:
A;Gene: ssm4; SPAC27D7.13c
A;Map position: 1

  Query Match             3.7%;  Score 105;  DB 2;  Length 670;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   80;  Conservative   58;  Mismatches  149;  Indels   86;  Gaps   15;

Qy     107 KKESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYS 166
           |: |  : |     : :::|  :      :::| | |    |: | |        :|| :
Db      73 KRPSVVKSR----KKGSENISNFMEKTKAIKQKSRREPSKFERSLARPLCITPIDSSTPT 128

Qy     167 RSSVDVLYTFANCSGLD-LIFGLNALLRTADLQWNSSNAQLL-LDYCSSKGYNISWELGN 224
           :::    || :    || | |    |       ||  ::|  ||  |       |:: |
Db     129 KTA--TFYTSSTTENLDELNFSTEELSSFDTTLLNSDTSKLSGLDDSSFMEEEFVWQVDN 186

Qy     225 ------------EPNSFLKK-----------ADIFINGSQLGEDYIQLHKLLRK----ST 257
                       |:||:            ::    : | |  :|:| | |     |
Db     187 VLQECEKKFTPHSKGSYLKENLKSELRKGRLDELMCENTALKEKIDKLNKELEKVEPQLT 246

Qy     258 FKNAKLYGPDVGQPRR-KTAKMLKSFLKAGGEV------------IDSVTWHHYYLNGRT 304
           |  :|     : :||  :  | || ||    |:            | :  :    || :|
Db     247 FLRSK---NSIEKPRNFRREKFLKKFLAMQKEIKYLRKRKLQIRKIPNYKYSDRSLNSKT 303

Qy     305 ATREDFLNPDVLD---IFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAA 361
              :|     |     : :| | || |:       |:| :  |
Db     304 PKSQDNWTTQVTPSSLLGVSEVSKVLQL-------KQVQVDITE---------------- 340

Qy     362 GFMWLDKLGLSARMGIEVVMRQVFFGAGN----YHLVDENFDPLPDYWLSLLFKKLVGTK 417
           : : |    | :: |  | | :    |:    : | :|||      :| | ::::|:  |
Db     341 -LVKIPKNPFSEKLTISNVNRYLNIVPGSLDLQFSLTNENF----VHWNSTVYQELLNLK 395

Qy     418 VLMASVQGSKRRK 430
              :|| | | |:
Db     396 SNNSSVDGVKTRR 408


RESULT 14
S00652
phosphoribosylamine-glycine ligase (EC 6.3.4.13) - fission yeast (Schizosaccharom
N;Alternate names: AIRSase; aminoimidazole ribotide synthetase; GARSase; glycinam
N;Contains: phosphoribosylamine-glycine ligase (EC 6.3.4.13); phosphoribosylformy
C;Species: Schizosaccharomyces pombe
C;Date: 07-Sep-1990 #sequence_revision 28-Oct-1994 #text_change 03-Jun-2002
C;Accession: S00652; T40496; T40422
R;McKenzie, R.; Schuchert, P.; Kilbey, B.
Curr. Genet. 12, 591-597, 1987
A;Title: Sequence of the bifunctional adel gene in the purine biosynthetic pathw
A;Reference number: S00652; MUID:89003164; PMID:3502942
A;Accession: S00652
A;Molecule type: DNA
A;Residues: 1-788 <MCK>
A;Cross-references: EMBL:X06601; NID:g4903; PIDN:CAA29820.1; PID:g4904
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoef
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21910
A;Accession: T40496
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-788 <WOO>
A;Cross-references: EMBL:AL021730; PIDN:CAA16823.1; GSPDB:GN00067; SPDB:SPBC4C3.0
A;Experimental source: strain 972h-; cosmid c4C3
R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21928
A;Accession: T40422
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 604-788 <SEE>
A;Cross-references: EMBL:AL035655; PIDN:CAB38600.1; GSPDB:GN00067; SPDB:SPBC405.0
A;Experimental source: strain 972h-; cosmid c405
C;Genetics:
```

```
A;Gene: ADE1; SPDB:SPBC405.01
A;Map position: 2
C;Superfamily: Saccharomyces cerevisiae ADE5 multifunctional protein; phosphoribosylamin
C;Keywords: cyclo-ligase; purine nucleotide biosynthesis
F;5-425/Domain: phosphoribosylamine-glycine ligase homology <PGL>
F;439-767/Domain: phosphoribosylformylglycinamidine cyclo-ligase homology <PFCL>

  Query Match             3.7%;  Score 104.5;  DB 1;  Length 788;
  Best Local Similarity   27.7%;  Pred. No. 13;
  Matches   70;  Conservative   36;  Mismatches  114;  Indels   33;  Gaps   11;

Qy    297 HYYLNGRTATRE--DFLNPDV-LDIFISSVQKVFQVVEST-RPGKKVWLGETSSAY---- 348
          |: || :  |||   : |  | :|     ||::  :|:|| |||    :|    :
Db    424 HHALNPKRKTREILTYENSGVSVDNGNEFVQRIKDLVKSTRRPGADADIGGFGGIFDLKQ 483

Qy    349 -GGGAPLL-SDTFAAGFMWLDKLGLSAR--MGIEVVMRQVFFGAGNYHLVDENFDPL--P 402
           |   ||| | |   |   |  | |:    :||::|   |        || :  :||
Db    484 AGWNDPLLVSATDGVGSKLLIALSLNKHDTVGIDLVAMNV------NDLVVQGAEPLIFL 537

Qy    403 DYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTK 462
          ||: :      | | :   |:| |:    :    |:    | :|       ::  |::
Db    538 DYFATGSLDLKVSTSFVEGVVKGCKQAGCALVGGETSEMPGLYHDGHYDANGTSVGAVSR 597

Qy    463 YLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTL--KMVD----DQTLPPLMEKPL 516
             || | |  : |  |       || |  |  || :|  |:|:    : |     :| :
Db    598 DDILPKPESFSKGDILL-------GLASDGVHSNGYSLVRKIVEYSDLEYTSVCPWDKNV 650

Qy    517 RPGSSLGLPAFSY 529
          | | || :|   |
Db    651 RLGDSLLIPTRIY 663


RESULT 15
F70411
adenylosuccinate synthetase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 16-Jul-1999
C;Accession: F70411
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: F70411
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-432 <AQF>
A;Cross-references: GB:AE000733; NID:g2983720; PIDN:AAC07286.1; PID:g2983724; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: purA
C;Superfamily: adenylosuccinate synthase

  Query Match             3.7%;  Score 104;  DB 2;  Length 432;
  Best Local Similarity   23.9%;  Pred. No. 5.8;
  Matches   96;  Conservative   39;  Mismatches  128;  Indels  138;  Gaps   22;

Qy     15 LLLLGPLGPLSPGALPRPAQAQDVVDLD------------------FFTQEPLHLVSPS 55
          :| | | | |        ||   ||||:                    |  :  ||| |
Db     51 ILHLLPTGILHEHVKGVIAQGM-VVDLEVLHKEVKNLEEKGIYVKERLFISDRAHLVMPY 109

Qy     56 FLSVTIDANLATDPRFLILLGSPK--LRTLARGLSPAYL-RFGGTKTDFLIFDPKKESTF 112
                  : |  |   |  : |  ||: |||: ::|  :    | | | |  |
Db    110 H------------KLLDSLFEKKKGIGTTLRGIGPAYMFKYG--RKGIRISDLKDEKRF 154

Qy    113 EERSYWQSQVNQDICKYGSIPPDVEEK-------LRLEWPYQEQLLLREHYQKKFKNSTY 165
             |   : | |  |       :: ||      | :  |:|||     | ::|| :
Db    155 ----YTLLEDNLDFVK------NICEKVFCEKFDLDINQIYEEQL----RYFEEFKENV- 199

Qy    166 SRSSVDVLYTFANCSGLDLIFGLNALLRTADL----QWNSSNAQLLLDYCSSKGYNISWE 221
             ||:|  |    |  | |    |   |:       ||||  |          :|
Db    200 ----VDLLRFFNTQKGSVLFEGAQGTLLDVDMGTYPYVTSSNASAL---------GLSNG 246

Qy    222 LGNEPNSFLKKADIFING--------------SQL-GEDYIQLHKLLRKSTFKNAKLYG 265
           |  |  |   :| |  |               ::| ||:  :| :|           ||
Db    247 TGMPPKYF---SDAFFLGVAKAYTTRVGEGPFPTELKGEEGEKLREL--------GGEYG 295

Qy    266 PDVGQPRR---KTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNP--------- 313
             |:|||        ||  ::  |            |:|   |: | |:
Db    296 STTGRPRRCGWLDLVALKYAVQVNG-----------LDGFVITKLDVLDTFDEVKVCVA 343

Qy    314 -----DVLDIFISSVQKVFQV--VESTRPGKKVWLGETSSA 347
                :|:| | :|  :: :|  |  |  |   |  |  |
Db    344 YELDGEVIDYFPASYSELIRVKPVYKTLKG---WKKSTKGA 381


Search completed: November 20, 2002, 11:38:11
Job time : 25 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 20, 2002, 11:34:27 ; Search time 14 Seconds
(without alignments)
1608.689 Million cell updates/sec

Title: US-09-759-207-2
Perfect score: 2842
Sequence: 1 MLLRSKPALPPPLMLLLLGP.....LPAFSYSFFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	112.5	4.0	356	1	Y670_METJA	Q58084 methanococc
2	111	3.9	688	1	YB9F_YEAST	P38338 saccharomyc
3	105.5	3.7	1031	1	TERT_EUPAE	O00939 euplotes ae
4	104.5	3.7	788	1	PUR2_SCHPO	P20772 s bifunctio
5	104	3.7	432	1	PURA_AQUAE	O67321 aquifex aeo
6	104	3.7	716	1	RRP2_IAKIT	O91742 influenza a
7	103	3.6	796	1	PHK_CLOAB	Q97je3 clostridium
8	100.5	3.5	805	1	GYRB_CHLPN	Q9z8r3 chlamydia p
9	99.5	3.5	837	1	XYNZ_CLOTM	P10478 clostridium
10	99	3.5	897	1	DYHC_HUMAN	Q14204 homo sapien
11	99	3.5	4644	1	DYHC_RAT	P38650 rattus norv
12	98.5	3.5	629	1	T3MH_HAEIN	P71366 haemophilus
13	98.5	3.5	654	1	NU5M_RHIST	P50367 rhizopus st
14	98.5	3.5	746	1	RHTA_RHIME	Q9z3q5 rhizobium m
15	98	3.4	716	1	RRP2_IAVI7	P31343 influenza a
16	97.5	3.4	454	1	YUA6_CAEEL	P52715 caenorhabdi
17	97.5	3.4	804	1	GYRB_CHLTR	O84193 chlamydia t
18	97.5	3.4	1314	1	SS22_YEAST	P25390 saccharomyc
19	96.5	3.4	595	1	THIC_BACHD	Q9kbj4 bacillus ha
20	96	3.4	327	1	XYNA_ASPAC	O59859 aspergillus
21	96	3.4	557	1	COX1_NEUCR	P03945 neurospora
22	96	3.4	716	1	RRP2_IAZI1	P13175 influenza a
23	96	3.4	1044	1	ITAV_MOUSE	P43406 mus musculu
24	95.5	3.4	358	1	VAL1_BCTV	P14991 beet curly
25	95.5	3.4	620	1	HEMA_MEASY	P28081 measles vir
26	95.5	3.4	5255	1	BACA_BACLI	O68006 b bacitraci
27	95	3.3	716	1	RRP2_IAKOR	P13170 influenza a
28	95	3.3	772	1	LP1G_DROME	P11997 drosophila
29	95	3.3	2214	1	SORL_HUMAN	Q92673 h sortilin-
30	95	3.3	4644	1	DYHC_MOUSE	Q9jhu4 mus musculu
31	94.5	3.3	592	1	INV2_DAUCA	Q39692 daucus caro
32	94	3.3	437	1	ERF1_XENLA	P35615 xenopus lae
33	94	3.3	804	1	GYRB_CHLMU	Q9pkk3 chlamydia m
34	93.5	3.3	657	1	GRAD_TREPA	O83062 treponema p
35	93.5	3.3	766	1	GAP1_SCHPO	P33277 schizosacch
36	93.5	3.3	1787	1	UVRA_CHLMU	Q9pk60 chlamydia m
37	93	3.3	449	1	A4M1_MOUSE	Q9jkc7 mus musculu
38	93	3.3	726	1	CATA_ECOLI	P13029 escherichia
39	93	3.3	2733	1	RRPB_CVMA5	P16342 murine coro
40	92.5	3.3	455	1	MURF_BUCAI	P57315 buchnera ap
41	92.5	3.3	738	1	YAS9_SCHPO	Q10145 schizosacch
42	92.5	3.3	828	1	YFA4_YEAST	P43585 saccharomyc
43	92	3.2	485	1	YA64_HAEIN	P71367 haemophilus
44	92	3.2	716	1	RRP2_IAANN	P21427 influenza a
45	92	3.2	1536	1	GLSF_ANTSP	Q06434 antithamnio

ALIGNMENTS

RESULT 1
Y670_METJA
ID Y670_METJA STANDARD; PRT; 356 AA.
AC Q58084;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0670.
GN MJ0670.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.";
RL Science 273:1058-1073(1996).
CC ---

CC ---
DR EMBL; U67514; AAB98664.1; -.
DR TIGR; MJ0670; -.
DR InterPro; IPR000051; SAM_bind.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 356 AA; 41683 MW; D7B8BA2E16A92E11 CRC64;

Query Match 4.0%; Score 112.5; DB 1; Length 356;
Best Local Similarity 21.2%; Pred. No. 0.38;
Matches 85; Conservative 48; Mismatches 143; Indels 125; Gaps 18;

```
Qy    126 ICKY-----GSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVD--------- 171
          | ||     |    |::|:|  |       |  ||   : :: ||:  : |
Db     14 IRKYMKIYNGKNEKDIKERLIKE-------LKEEHVLVETEDGTYTLKAEDEEEMMHSKV 66

Qy    172 -----VLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEP 226
               :| ||  |          :  ||    || : :|| ||  |||    |
Db     67 GALKEAIYKFAKPS-----------KITDL----SNPR-VLDLCSGMGYNAIAALHYNK 109

Qy    227 NS---------------FL-----------KKADIFINGSQLGEDYIQLHKLLRKSTF 258
```

```
Db    110 NAEIDMVEICEEVLFLTLFLDIPYKEHEIIKDKVREYFLN--KIGIEY--------KSDY 159

Qy    259 KNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTAT--REDFLNPDVL 316

Db    160 DNINLY---VGDARKFIIKSDKKY---------NVVFHDAFSPKRDPTLYTYDFL----- 202

Qy    317 DIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMG 376

Db    203 -------KEIYKRMEDN--GVLI-------SYSSAIPFRSALVDCGFVISEKESVGRKRG 246

Qy    377 IEVVMRQVFFGAGNYHLVDENFD-------PLPDYWLSLLFKKLVGTKVLMASVQGSKRR 429

Db    247 ITLAYKNPNFKPNRINEVDERVIALSVIALPYRDETLSLTKDKIIEDREERREKLKEKLI 306

Qy    430 KLRVYLHCTNTDNPRYKEGDLTLYA--INLHNVTKYLRLPY 468

Db    307 KIGKYLSTKQIKKGNIPEEILKIQKEDLNSSEIIKKMRLKF 347


RESULT 2
YB9F_YEAST
ID   YB9F_YEAST     STANDARD;      PRT;   688 AA.
AC   P38338;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Hypothetical 80.4 kDa protein in POP4-SHM1 intergenic region.
GN   YBR259W OR YBR1727.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=93220397; PubMed=8465606;
RA   Doignon F., Biteau N., Crouzet M., Aigle M.;
RT   "The complete sequence of a 19,482 bp segment located on the right
RT   arm of chromosome II from Saccharomyces cerevisiae.";
RL   Yeast 9:189-199(1993).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70529; CAA49923.1; -.
DR   EMBL; Z36128; CAA85222.1; -.
DR   PIR; S32961; S32961.
DR   SGD; S0000463; YBR259W.
KW   Hypothetical protein.
SQ   SEQUENCE   688 AA;  80426 MW;  0BA84837BD7A4B30 CRC64;

  Query Match             3.9%;  Score 111;  DB 1;  Length 688;
  Best Local Similarity   22.5%;  Pred. No. 1.3;
  Matches   67;  Conservative   45;  Mismatches   94;  Indels   92;  Gaps   16;

Qy    126 ICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSS----------VDVLYT 175

Db    164 MAEYSSWKWDSDDKRQLQFMYEFRMKLKECLVKFYENFDLQKSSDPLKELIIPWEKIVYV 223

Qy    176 FANCSGLDLIFGLNALLRTADLQWNSSN--------AQLLLD-----YCSSKGY------ 216

Db    224 -ANC--IDAFTGEQVRIDGAELIWTSKNLVFSSISSAVLRLNDLQNMFSAFRPYGEEALV 280

Qy    217 -------NISWELGNEPNSFLKKA---DIF--INGSQLG--EDYIQLHKLLRK------- 255

Db    281 QDFAHIRSLKWDSNDKVESLIRALIFNDMFPYFNKEQVDTKADGIFFLRLLRKNFKEHIN 340

Qy    256 -------------STFKN--AKLYGPDVGQPRRKTAKMLKSFLKAGGEV------IDSV 293

Db    341 DVKDFHIQVIKYLNSQFKNNYSTLMTSSKTQDRRKSHNMPSSILDDGNKIGMHVSPIDE- 399

Qy    294 TWHHYYLNG----------RTATREDFLNPDVLDIFISSVQKVFQVVESTR---PGKK 338

Db    400 -YSHFIDNDEPLWRDKVYPKIYTNEQTPTPDASAIFDS--HKIYAIISLLRYYLPEKR 454


RESULT 3
TERT_EUPAE
ID   TERT_EUPAE     STANDARD;      PRT;  1031 AA.
AC   O00939;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic
DE   subunit) (Telomerase subunit P123).
OS   Euplotes aediculatus.
OC   Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
OC   Euplotida; Euplotidae; Euplotes.
OX   NCBI_TaxID=5940;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97274210; PubMed=9110970;
RA   Lingner J., Hughes T.R., Shevchenko A., Mann M., Lundblad V.,
RA   Cech T.R.;
RT   "Reverse transcriptase motifs in the catalytic subunit of
RT   telomerase.";
RL   Science 276:561-567(1997).
CC   -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR
CC       THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT
CC       ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS
CC       SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE
CC       SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.
CC       TELOMERASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U95964; AAC47515.1; -.
DR   InterPro; IPR000477; RVTse.
DR   InterPro; IPR003545; Telomerase_RT.
DR   Pfam; PF00078; rvt; 1.
DR   PRINTS; PR01365; TELOMERASERT.
KW   Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;
KW   DNA-binding.
SQ   SEQUENCE   1031 AA;  122562 MW;  57B87A63A1FED60F CRC64;

  Query Match             3.7%;  Score 105.5;  DB 1;  Length 1031;
  Best Local Similarity   20.9%;  Pred. No. 6;
  Matches   81;  Conservative   56;  Mismatches  133;  Indels  117;  Gaps   17;

Qy    102 LIFDPKKESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFK 161

Db    694 LIVEAKQRNYFKKDNLLQPVIN--ICQYNYI----------NFNGKFY 729

Qy    162 NST------YSRSSVDVLYTFANCSGLDLIF-----------GLNALLRTAD----LQWN 200

Db    730 KQTKGIPQGLCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLITTQ 789

Qy    201 SSNAQLLLDYCSSKGYNISWELGNE-------------PNSFLKKADIFINGSQLGEDYI 247

Db    790 ENNAVLFIE----KLINVSRENGFKFNMKKLQTSFPLSPSKFAKYGMDSVEEQNIVQDYC 845

Qy    248 QL------HKLLRKSTFKNAKLYG----PDVGQPRRKTA----KMLKSFLKAGGEVIDSV 293
```

```
Db    846 DWIGISIDMKTLALMPNINLRIEGILCTLNLNMQTKKASMWLKKKLKSFL------MNNI 899

Qy    294 TWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAP 353
          |  ||:   :| | ||| |  :  :|||   |  |  :  :   |  |  :|      :
Db    900 T--HYF--RKTITTEDFANKTLNKLFISGGYKYMQCAKEYKDHFKKNLAMSSMIDLEVSK 955

Qy    354 LLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLL---- 409
          ::     | |:|              ::   ||  :|          ||::|| |
Db    956 IIYSVTRAFFKYL----------VCNIKDTIFGEEHY----------PDFFLSTLKHFI 994

Qy    410 ----FKKLVGTKVLMASVQGSKRRKLR 432
              || :  :| |  :  :|  ||:
Db    995 EIFSTKKYIFNRVCM--ILKAKEAKLK 1019
```

```
RESULT 4
PUR2_SCHPO
ID   PUR2_SCHPO     STANDARD;      PRT;   788 AA.
AC   P20772; Q9UUM5;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Bifunctional purine biosynthetic protein ADE1 [Includes:
DE   Phosphoribosylamine--glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide
DE   ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase);
DE   Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1) (AIRS)
DE   (Phosphoribosyl-aminoimidazole synthetase) (AIR synthase)].
GN   ADE1 OR SPBC4C3.02C OR SPBC405.01.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=89003164; PubMed=3502942;
RA   McKenzie R., Schuchert P., Kilbey B.;
RT   "Sequence of the bifunctional ade1 gene in the purine biosynthetic
RT   pathway of the fission yeast Schizosaccharomyces pombe.";
RL   Curr. Genet. 12:591-597(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- CATALYTIC ACTIVITY: ATP + 5-phospho-D-ribosylamine + glycine = ADP
CC       + phosphate + N(1)-(5-phospho-D-ribosyl)glycinamide.
CC   -!- CATALYTIC ACTIVITY: ATP + 2-(formamido)-N(1)-(5-phospho-D-
CC       ribosyl)acetamidine = ADP + phosphate + 5-amino-1-(5-phospho-D-
CC       ribosyl)imidazole.
CC   -!- PATHWAY: De novo purine biosynthesis; second step.
CC   -!- PATHWAY: De novo purine biosynthesis; fifth step.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GARS FAMILY.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE AIRS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X06601; CAA29820.1; -.
DR   EMBL; AL021730; CAA16823.1; -.
DR   EMBL; AL035655; CAB38600.1; -.
DR   PIR; S00652; S00652.
DR   HSSP; P08178; 1CLI.
DR   InterPro; IPR000728; AIRS_related.
DR   InterPro; IPR000115; Gars.
DR   InterPro; IPR004733; PurM_cligase.
DR   Pfam; PF00586; AIRS; 1.
DR   Pfam; PF01071; GARS; 1.
DR   Pfam; PF02769; AIRS_C; 1.
DR   Pfam; PF02842; GARS_B; 1.
DR   Pfam; PF02843; GARS_C; 1.
DR   Pfam; PF02844; GARS_N; 1.
DR   TIGRFAMs; TIGR00877; purD; 1.
DR   TIGRFAMs; TIGR00878; purM; 1.
DR   PROSITE; PS00184; GARS; 1.
KW   Multifunctional enzyme; Purine biosynthesis; Ligase.
FT   DOMAIN        1    430       GARS.
FT   DOMAIN      440    750       AIRS.
SQ   SEQUENCE   788 AA;  85231 MW;  0FDE64EEA5F9095D CRC64;

  Query Match             3.7%;  Score 104.5;  DB 1;  Length 788;
  Best Local Similarity   27.7%;  Pred. No. 4.8;
  Matches   70;  Conservative   36;  Mismatches  114;  Indels   33;  Gaps   11;

Qy    297 HYYLNGRTATRE--DFLNPDV-LDIFISSVQKVFQVVEST-RPGKKVWLGETSSAY---- 348
          |: || :  |||   : |  | :|     ||::  :|:|| |||    :|     :
Db    424 HHALNPKRKTREILTYENSGVSVDNGNEFVQRIKDLVKSTRRPGADADIGGFGGIFDLKQ 483

Qy    349 -GGGAPLL-SDTFAAGFMWLDKLGLSAR--MGIEVVMRQVFFGAGNYHLVDENFDPL--P 402
           |   ||| | |   |   |  | |:    :||::|   |        || :  :||
Db    484 AGWNDPLLVSATDGVGSKLLIALSLNKHDTVGIDLVAMNV------NDLVVQGAEPLIFL 537

Qy    403 DYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTK 462
          ||: :      | |  :   |:| |:    :    |:     | :|       ::  |::
Db    538 DYFATGSLDLKVSTSFVEGVVKGCKQAGCALVGGETSEMPGLYHDGHYDANGTSVGAVSR 597

Qy    463 YLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTL--KMVD----DQTLPPLMEKPL 516
            || | |  : |  |       || |  |  || :|  |:|:    : |     :| :
Db    598 DDILPKPESFSKGDILL-------GLASDGVHSNGYSLVRKIVEYSDLEYTSVCPWDKNV 650

Qy    517 RPGSSLGLPAFSY 529
          | | || :|   |
Db    651 RLGDSLLIPTRIY 663


RESULT 5
PURA_AQUAE
ID   PURA_AQUAE     STANDARD;      PRT;   432 AA.
AC   O67321;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
DE   (AdSS) (AMPSase).
```

```
GN   PURA OR AQ_1290.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC   Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
CC   -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DE NOVO PATHWAY OF PURINE
CC       NUCLEOTIDE BIOSYNTHESIS.
CC   -!- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
CC       adenylosuccinate.
CC   -!- PATHWAY: AMP biosynthesis; first committed step.
CC   -!- SIMILARITY: BELONGS TO THE ADENYLOSUCCINATE SYNTHETASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000733; AAC07286.1; -.
DR   HSSP; P12283; 1ADE.
DR   InterPro; IPR001114; Asucc_synthtase.
DR   Pfam; PF00709; Adenylsucc_synt; 1.
DR   ProDom; PD001188; Asucc_synthtase; 1.
DR   TIGRFAMs; TIGR00184; purA; 1.
DR   PROSITE; PS01266; ADENYLOSUCCIN_SYN_1; 1.
DR   PROSITE; PS00513; ADENYLOSUCCIN_SYN_2; 1.
KW   Purine biosynthesis; Ligase; GTP-binding; Complete proteome.
FT   NP_BIND      12     18       GTP (POTENTIAL).
FT   ACT_SITE    137    137       BY SIMILARITY.
FT   ACT_SITE    144    144       BY SIMILARITY.
SQ   SEQUENCE   432 AA;  48854 MW;  ACA52CE913DE62E7 CRC64;

  Query Match             3.7%;  Score 104;  DB 1;  Length 432;
  Best Local Similarity   23.9%;  Pred. No. 2.2;
  Matches   96;  Conservative   39;  Mismatches  128;  Indels  138;  Gaps   22;

Qy      15 LLLLGPLGPLSPGALPRPAQAQDVVDLD------------------FFTQEPLHLVSPS 55
           :| | | | |        ||   ||||:                    |  :  ||| |
Db      51 ILHLLPTGILHEHVKGVIAQGM-VVDLEVLHKEVKNLEEKGIYVKERLFISDRAHLVMPY 109

Qy      56 FLSVTIDANLATDPRFLILLGSPK--LRTLARGLSPAYL-RFGGTKTDFLIFDPKKESTF 112
                  : |  |   |  : |  ||: |||: ::|  :    | | | |  |
Db     110 H-------------KLLDSLFEKKKGIGTTLRGIGPAYMFKYG--RKGIRISDLKDEKRF 154

Qy     113 EERSYWQSQVNQDICKYGSIPPDVEEK-------LRLEWPYQEQLLLREHYQKKFKNSTY 165
           |   : | |  |       :: ||       | :   |:|||     | ::|| :
Db     155 ----YTLLEDNLDFVK------NICEKVFCEKFDLDINQIYEEQL----RYFEEFKENV- 199

Qy     166 SRSSVDVLYTFANCSGLDLIFGLNALLRTADL----QWNSSNAQLLLDYCSSKGYNISWE 221
           ||:|  |    |  |  |    |   |:       ||||  |          :|
Db     200 ----VDLLRFFNTQKGSVLFEGAQGTLLDVDMGTYPYVTSSNASAL---------GLSNG 246

Qy     222 LGNEPNSFLKKADIFING---------------SQL-GEDYIQLHKLLRKSTFKNAKLYG 265
           |  |  |   :| |  |                ::| ||:  :| :|          ||
Db     247 TGMPPKYF---SDAFFLGVAKAYTTRVGEGPFPTELKGEEGEKLREL--------GGEYG 295

Qy     266 PDVGQPRR---KTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNP--------- 313
           |:|||       ||  ::  |           |:|   |: | |:
Db     296 STTGRPRRCGWLDLVALKYAVQVNG-----------LDGFVITKLDVLDTFDEVKVCVA 343

Qy     314 -----DVLDIFISSVQKVFQV--VESTRPGKKVWLGETSSA 347
                :|:| | :|  :: :|  |  | |  |   |  |
Db     344 YELDGEVIDYFPASYSELIRVKPVYKTLKG---WKKSTKGA 381


RESULT 6
RRP2_IAKIT
ID   RRP2_IAKIT     STANDARD;      PRT;   716 AA.
AC   O91742;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   RNA-directed RNA polymerase subunit P2 (EC 2.7.7.48) (Polymerase
DE   acidic protein) (PA).
OS   Influenza A virus (strain A/Kitakyushu/159/93).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=62478;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98406205; PubMed=9733841;
RA   Lindstrom S.E., Hiromoto Y., Nerome R., Omoe K., Sugita S.,
RA   Yamazaki Y., Takahashi T., Nerome K.;
RT   "Phylogenetic analysis of the entire genome of influenza A (H3N2)
RT   viruses from Japan: evidence for genetic reassortment of the six
RT   internal genes.";
RL   J. Virol. 72:8021-8031(1998).
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: INFLUENZA RNA POLYMERASE IS COMPOSED OF THREE SUBUNITS:
CC       P1 (OR PB1), P2 (OR PA), AND P3 (OR PB2).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA VIRUSES POLYMERASE PA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF037424; AAC63455.1; -.
DR   InterPro; IPR001009; RNA_pol_P2.
DR   Pfam; PF00603; Flu_PA; 1.
KW   Transferase; RNA-directed RNA polymerase.
SQ   SEQUENCE   716 AA;  82749 MW;  51A8DF9A74AB5159 CRC64;

  Query Match             3.7%;  Score 104;  DB 1;  Length 716;
  Best Local Similarity   19.7%;  Pred. No. 4.5;
  Matches  102;  Conservative   78;  Mismatches  194;  Indels  144;  Gaps   25;

Qy      57 LSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERS 116
           ::  |:  | | || : |   |          | | |     : ||: | | |  :
Db     253 VNAKIEPFLKTTPRPIKLPNGP----------PCYQR-----SKFLLMDALKLSIED--- 294

Qy     117 YWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTF 176
               |   : |  | :|     : :|  : ::|  ::: |  :|  || | |   ||
Db     295 --PSHEGEGIPLYDAI-----KCIRTFFGWKEPYIVKPH--EKGINSNYLLSWKQVLAEL 345

Qy     177 ANCSGLDLIFGLNALLRTADLQW----NSSNAQLLLDYCSS----KGYN---------IS 219
           :    : |     : :|: |:|    | :  ::  | |       | |:         |
Db     346 QDIETEEKIPRTKNMKKTSQLKWALGENMAPEKVDFDNCRDISDLKQYDSDEPELRSLSS 405

Qy     220 WELGNEPNSFLKKAD-IFINGSQLGEDYIQLHKL--LRKSTFKNAKLYGPDVGQPRRKTA 276
           | : || |   :  | |:|   ::|||   :  :  :|:: |                ||
Db     406 W-IQNEFNKACELTDSIWIELDEIGEDVAPIEYIASMRRNYF---------------TA 448

Qy     277 KMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPG 336
           ::  |  :|   ::  |  :   ||   |  :||         |  : |       |: |
Db     449 EV--SHCRATEYIMKGVYINTALLNASCAAMDDFQ-------LIPMISKC-----RTKEG 494

Qy     337 KKVWLGETSSAYG----GGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYH 392
           ::       :: ||    | : | :||    |: ::                    :
Db     495 RR-----KTNLYGFIIKGRSHLRNDTDVVNFVSME----------------------FS 526
```

```
Qy     393 LVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTL 452
           | |   :  | |      ::    :|: |  |   | : :|:   .|   : | |
Db     527 LTDPRLE--PHKWEKYCVLEI--GDMLLRSAIGQMSRPMFLYVRTNGTSKIKMKWGMEMR 582

Qy     453 YAI--NLHNVTKYLRLPYPFSNKQVDKYLLR------PLG--PHG-----------LLS 490
             :  :|  :   :        | : |          |:|  | |            ||:
Db     583 RCLLQSLQQIESMIEAESSVKEKDMTKEFFENKSEAWPIGESPKGVEEGSIGKVCRTLLA 642

Qy     491 KSV--------QLNGLTLKMVDDQTLPPLMEKPLRPGS 520
           |||        || | : :      :   :   | ||:
Db     643 KSVFNSLYASPQLEGFSAESRKLLLVVQALRDNLEPGT 680


RESULT 7
PHK_CLOAB
ID   PHK_CLOAB      STANDARD;      PRT;   796 AA.
AC   Q97JE3;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable phosphoketolase (EC 4.1.2.-).
GN   CAC1343.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
CC   -!- COFACTOR: Thiamine pyrophosphate (Potential).
CC   -!- SIMILARITY: BELONGS TO THE XFP FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE007645; AAK79311.1; -.
DR   PROSITE; PS60002; PHOSPHOKETOLASE_1; 1.
DR   PROSITE; PS60003; PHOSPHOKETOLASE_2; 1.
DR   PROSITE; PS00187; TPP_ENZYMES; FALSE_NEG.
KW   Lyase; Flavoprotein; Thiamine pyrophosphate; Complete proteome.
SQ   SEQUENCE   796 AA;  90640 MW;  4B8219DC9778FAEF CRC64;

  Query Match             3.6%;  Score 103;  DB 1;  Length 796;
  Best Local Similarity   21.6%;  Pred. No. 6.3;
  Matches   77;  Conservative   51;  Mismatches  114;  Indels  114;  Gaps   18;

Qy      27 GALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVTIDA---NLATDPRF-LILLGSPKLRT 82
           |  |    : :|   :     | | :|:   |::  :|   |  : |:: :|:| :||
Db     238 GWKPYFVEGEDPETMHKLMAETLDIVTEEILNIQKNARENNDCSRPKWPMIVLRTPK--- 294

Qy      83 LARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRL 142
                   | |   |:       :    | |:   ||   : :| :   |  |
Db     295 -----------GWTGPKFV------DGVPNEGSFRAHQVPLAVDRYHTENLDQLE---- 332

Qy     143 EW--PYQEQLLLREHYQ--KKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALLRTADLQ 198
           ||   |: : |  |:|:   : :  |   :         ||  ||     |   ||| | :
Db     333 EWLKSYKPEELFDENYRLIPELEELTPKGNKRMAANLHAN-GGL-----LLRELRTPDFR 386

Qy     199 WNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDYIQLHKLLR---- 254
                   ||                        |:   || : :| |:| | :|
Db     387 ----------DYA--------------------VDVPTPGSTVKQDMIELGKYVRDVVK 415

Qy     255 -KSTFKNAKLYGPD----------VGQPRRKTAKMLK---SFLKAGGEVIDSVTWHH-- 297
                :| :::|||            |  |:  ::: :    ||   | ::||:   |
Db     416 LNEDTRNFRIFGPDETMSNRLWAVFEGTKRQWLSEIKEPNDEFLSNDGRIVDSMLSEHLC 475

Qy     298 ------YYLNGRTATREDFLNPDVLDIFISSVQKVFQVVES--TRPGKKVWLGETS 345
                 | | ||              | :| :   ::|:|  |: ||  ||  ||
Db     476 EGWLEGYLLTGRHG-------------FFASYEAFLRIVDSMITQHGK--WLKVTS 516


RESULT 8
GYRB_CHLPN
ID   GYRB_CHLPN     STANDARD;      PRT;   805 AA.
AC   Q9Z8R3; Q9JQ44;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA gyrase subunit B (EC 5.99.1.3).
GN   GYRB OR CPN0275 OR CP0484.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg.J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC       STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC       INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC       DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC   -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC       of double-stranded DNA.
CC   -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC       BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC       ENZYME FORMS AN A2B2 TETRAMER.
CC   -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE001612; AAD18424.1; -.
```

```
DR   EMBL; AE002210; AAF38314.1; -.
DR   EMBL; AP002546; BAA98485.1; -.
DR   HSSP; P06982; 1AJ6.
DR   PHCI-2DPAGE; Q9Z8R3; -.
DR   TIGR; CP0484; -.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR002288; DNA_gyraseB_C.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002936; DNAprim_toprim.
DR   Pfam; PF00204; DNA_gyraseB; 1.
DR   Pfam; PF00986; DNA_gyraseB_C; 1.
DR   Pfam; PF01751; Toprim; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00418; TPI2FAMILY.
DR   ProDom; PD000616; DNA_topoisoII; 1.
DR   ProDom; PD149633; DNA_gyraseB_C; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   TIGRFAMs; TIGR01059; gyrB; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Topoisomerase; Isomerase; ATP-binding; Complete proteome.
SQ   SEQUENCE   805 AA;  90571 MW;  C082DF4CCEC71ECC CRC64;

  Query Match             3.5%;  Score 100.5;  DB 1;  Length 805;
  Best Local Similarity   20.4%;  Pred. No. 9.9;
  Matches  121;  Conservative   71;  Mismatches  169;  Indels  231;  Gaps   32;

Qy      48 PLHLVSPS---------------FLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYL 92
           ||  || |               | : | |       | :::    :|  | ||::
Db     161 PLQYVSVSDRQGTEIVFYPDPKIFSTCTFD-------RSILMKRLRELAFLNRGIT---- 209

Qy      93 RFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLL 152
                 ::|:  :: :|:: :::           | |   |
Db     210 ---------IVFEDDRDVSFDKVTFFYE---------GGIQSFVS-------------- 236

Qy     153 REHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCS 212
             |  : | | :|    : :|      | |       :   | |||||  ::|:  | :
Db     237 ---YLNQNKESLFS----EPIYICGTRVGDD-----GEIEFEAALQWNSGYSELVYSYAN 284

Qy     213 SKGYNI-SWELGNEPNSFLKKADIFINGSQLGEDYIQLHKLLRKSTFKNAK--LYGPDVG 269
           || : : |     |       ||       ||: | |      || |  | | |:
Db     285 ----NIPTRQGGTHLTGFSTALTRVIN------TYIKAHNLA-----KNNKLALTGEDI- 328

Qy     270 QPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDV----------LDI 318
                     : |   : ||   :    |:  |::   | ||           | |
Db     329 -------------REGLTAVISVKVPNPQFEGQ--TKQKLGNSDVSSVAQQVVGEALTI 372

Qy     319 FISS--------VQKVF-----------------QVVESTR-PGK-----------K 338
           |          | |||                  ::| | |||           :
Db     373 FFEENPQIARMIVDKVFVAAQAREAAKKARELTLRKSALDSARLPGKLIDCLEKDPEKCE 432

Qy     339 VWLGETSSAYGGGA-----------PLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFF 386
           ::: |  || || |           |:          | |:  :  :|  :
Db     433 MYIVEGDSA-GGSAKQGRDRRFQAILPIRGKILNVEKARLQKIFQNQEIGTIIAALGCGI 491

Qy     387 GAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM--ASVQGSKRRKLRV---YLHCTN-- 439
           || |::|              | ::::    ::|  | | ||  | | :   | | |
Db     492 GADNFNL-----------SKLRYRRI----IIMTDADVDGSHIRTLLLTFFYRHMTALI 535

Qy     440 -------TDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPH--GLLS 490
                 | ||          |:|      |. | |::| |||  || :   :|
Db     536 ENECVYIAQPPLYK------------VSKKKDFRYILSEKEMDSYLLM-LGTNESSILF 581

Qy     491 KSV--QLNGLTLK-----MVDDQTLPPLMEKPLRPGS--------SLGLPAF 527
           ||   :| |  |:      ::| ::    :||   | |         :| | :
Db     582 KSTERELRGEALESFINVILDVESFINTLEKKAIPFSEFLEMYKEGIGYPLY 633


RESULT 9
XYNZ_CLOTM
ID   XYNZ_CLOTM     STANDARD;      PRT;   837 AA.
AC   P10478;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Endo-1,4-beta-xylanase Z precursor (EC 3.2.1.8) (Xylanase Z)
DE   (1,4-beta-D-xylan xylanohydrolase Z).
GN   XYNZ.
OS   Clostridium thermocellum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1515;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCIB 10682;
RX   MEDLINE=89008072; PubMed=3139632;
RA   Grepinet O., Chebrou M.-C., Beguin P.;
RT   "Nucleotide sequence and deletion analysis of the xylanase gene
RT   (xynZ) of Clostridium thermocellum.";
RL   J. Bacteriol. 170:4582-4588(1988).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS) OF 515-837.
RC   STRAIN=NCIB 10682;
RX   MEDLINE=95393242; PubMed=7664125;
RA   Dominguez R., Souchon H., Spinelli S., Dauter Z., Wilson K.S.,
RA   Chauvaux S., Beguin P., Alzari P.M.;
RT   "A common protein fold and similar active site in two distinct
RT   families of beta-glycanases.";
RL   Nat. Struct. Biol. 2:569-576(1995).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC       linkages in xylans.
CC   -!- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
CC       WELL AS 1N OTHER C.THERMOCELLUM CELLULOSOME ENZYMES. THIS DOMAIN
CC       MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC       HYDROLASES).
CC   -!- SIMILARITY: CONTAINS 1 XYNZ-TYPE CELLULOSE-BINDING DOMAIN (CBD).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22624; AAA23286.1; -.
DR   PIR; A31842; A31842.
DR   PDB; 1XYZ; 29-JAN-96.
DR   InterPro; IPR005084; CBM_6.
DR   InterPro; IPR002105; Dockerin_1.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR001000; Glyco_hydro_10.
DR   InterPro; IPR000379; Ser_estrs_site.
DR   Pfam; PF00331; Glyco_hydro_10; 1.
DR   Pfam; PF00404; Dockerin_1; 2.
DR   Pfam; PF03422; CBM_6; 1.
DR   PRINTS; PR00134; GLHYDRLASE10.
DR   PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR   PROSITE; PS00448; CLOS_CELLULOSOME_RPT; 2.
DR   PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.
KW   Xylan degradation; Hydrolase; Glycosidase; Repeat; Signal;
KW   3D-structure.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    837       ENDO-1,4-BETA-XYLANASE Z.
FT   ACT_SITE    645    645       PROTON-DONOR.
FT   ACT_SITE    754    754       NUCLEOPHILE.
FT   DOMAIN      328    416       CELLULOSE-BINDING (POTENTIAL).
FT   DOMAIN      430    487       2 X 24 AA APPROXIMATE REPEATS.
FT   REPEAT      430    453       1.
FT   REPEAT      464    487       2.
FT   DISULFID    783    789       BY SIMILARITY.
SQ   SEQUENCE   837 AA;  92262 MW;  DD4C29F04D12B6CD CRC64;

  Query Match             3.5%;  Score 99.5;  DB 1;  Length 837;
  Best Local Similarity   19.3%;  Pred. No. 12;
```

```
  Matches  63;  Conservative   48;  Mismatches  107;  Indels  109;  Gaps   16;

Qy     145 PYQEQLLLRE------HYQKKF-----KNSTYSRSSVDVLYTFANCSGLDLIFGLNALLR 193
           |    :| ||        : ||     : : :  |  | |  ||  :|:          :|
Db     543 PTYNSILQREFSMVVCENEMKFDALQPRQNVFDFSKGDQLLAFAERNGMQ--------MR 594

Qy     194 TADLQWNSSNAQLLLD------------------YCSSKGYNISWELGNE-----PNSF 229
             | |:: |   | :                        ||  : |:: ||       |
Db     595 GHTLIWHNQNPSWLTNGNWNRDSLLAVMKNHITTVMTHYKGKIVEWDVANECMDDSGNGL 654

Qy     230 LKKADIFINGSQLGEDYIQLHKLLRKSTFKNAKLYGP------------DVGQPRRKTAK 277
             :: |: |   :|:||:          |: |:   |             |:|
Db     655 --RSSIWRN--VIGQDYLDY-------AFRYAREADPDALLFYNDYNIEDLGPKSNAVFN 703

Qy     278 MLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDV-----LDIFISSVQKVFQVVES 332
           |:|| :|  |  || | :  :::|| :      :: ::     : : :|  :   :: :|
Db     704 MIKS-MKERGVPIDGVGFQCHFINGMSPEYLASIDQNIKRYAEIGVIVSFTEIDIRIPQS 762

Qy     333 TRPG------------KKVWLGETSSAYGGGAPLLSDTFAAGFMW--LDKLGLSARMGI 377
             |               |: |   :           :||    ||   ||
Db     763 ENPATAFQVQANNYKELMKICLANPN----------CNTFV---MWGFTDKY-------- 801

Qy     378 EVVMRQVFFGAGNYHLVDENFDPLPDY 404
             :   | | ||  : | |::| | |
Db     802 -TWIPGTFPGYGNPLIYDSNYNPKPAY 827


RESULT 10
DYHC_HUMAN
ID   DYHC_HUMAN     STANDARD;      PRT;   897 AA.
AC   Q14204; Q92814;
DT   01-NOV-1997 (Rel. 35, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Dynein heavy chain, cytosolic (DYHC) (Cytoplasmic dynein heavy chain
DE   1) (DHC1) (Fragment).
GN   DNCH1 OR DNECL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96234671; PubMed=8666668;
RA   Vaisberg E.A., Grissom P.M., McIntosh J.R.;
RT   "Mammalian cells express three distinct dynein heavy chains that are
RT   localized to different cytoplasmic organelles.";
RL   J. Cell Biol. 133:831-842(1996).
RN   [2]
RP   SEQUENCE OF 755-895 FROM N.A.
RX   MEDLINE=94043467; PubMed=8227145;
RA   Vaisberg E.A., Koonce M.P., McIntosh J.R.;
RT   "Cytoplasmic dynein plays a role in mammalian mitotic spindle
RT   formation.";
RL   J. Cell Biol. 123:849-858(1993).
CC   -!- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
CC       MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
CC       ORGANELLES ALONG MICROTUBULES.
CC   -!- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
CC       INTERMEDIATE AND LIGHT CHAINS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U53530; AAB09727.1; -.
DR   EMBL; L23958; AAA16065.1; -.
DR   Genew; HGNC:2961; DNCH1.
DR   MIM; 600112; -.
KW   Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT   NON_TER       1      1
FT   DOMAIN       42     74       COILED COIL (POTENTIAL).
FT   DOMAIN      102    123       COILED COIL (POTENTIAL).
FT   DOMAIN      228    244       COILED COIL (POTENTIAL).
FT   NP_BIND     777    784       ATP (POTENTIAL).
FT   CONFLICT    812    812       R -> M (IN REF. 2).
FT   NON_TER     897    897
SQ   SEQUENCE   897 AA;  103012 MW;  7A95514D06CA7D42 CRC64;

  Query Match             3.5%;  Score 99;  DB 1;  Length 897;
  Best Local Similarity   19.9%;  Pred. No. 15;
  Matches  114;  Conservative   81;  Mismatches  174;  Indels  204;  Gaps   27;

Qy      41 LDFFTQEPLHLVSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTD 100
           :|   ::|   | |  |   :|| |     |       :||  |       |  |  | :
Db     214 IDQMKEQPWVSVQPRKLRQNLDALLNQLKSF-----PARLRQYASYEFVQRLLKGYMKIN 268

Qy     101 FLIFDPKKESTFEERSYWQSQVNQ-----------------DICKYGSIPPDV------ 136
           |: : | |:  ::| :|:  : :                  |: |  :|  ||
Db     269 MLVIELKSEA-LKDR-HWKQLMKRLHVNWVVSELTLGQIWDVDLQKNEAIVKDVLLVAQG 326

Qy     137 ----EE---KLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLN 189
               ||  .::|  |   |  |:  :|| | :                       || | :
Db     327 EMALEEFLKQIREVWNTYELDLV--NYQNKCR---------------------LIRGWD 362

Qy     190 ALLRTADLQWNSSNAQLLLDYCSSKGYNI------SWELGNEPNSFLKKADIFIN----- 238
           |        || :|  |  |     | :      |||  :: |  :   |::|:
Db     363 DLFNKVKEHINSVSAMKLSPY-----YKVFEEDALSWE--DKLNRIMALFDVWIDVQRRW 415

Qy     239 --------GS--------------QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKT 275
                   ||                :  ::: | | : ||      |    | :   :
Db     416 VYLEGIFTGSADIKHLLPVETQRFQSISTEFLALMKKVSKSPLVMDVLNIQGVQRSLERL 475

Qy     276 AKMLKSFLKAGGEVI--DSVTWHHYYLNGRTATREDFLNPDVLDIF-----ISSVQKVFQ 328
           | :|    || || :  :  ::  :|  |          : |:|:|      :: :|| |
Db     476 ADLLGKIQKALGEYLERERSSFPRFYFVG---------DEDLLEIIGNSKNVAKLQKHF- 525

Qy     329 VVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGA 388
                   ||:: | :|        :|::  :        ||:|:| | ||: :
Db     526 --------KKMFAGVSSI-------ILNEDNSV------VLGISSREGEEVMFKTP---- 560

Qy     389 GNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEG 448
                 |    : ||:|: |::   :| :| :      :: ::   |: |   |
Db     561 -----VSITEHPKINEWLTLVEKEM---RVTLAKLLAESVTEVEIFGKATSIDPNTY--- 609

Qy     449 DLTLYAINLHNVTKYLRLPYPFSNKQVDKY------------------LLRPLGPHG-- 487
                      :|              :|||                    |  :|  |
Db     610 -----------IT------------WIDKYQAQLVVLSAQIAWSENVETALSSMGGGGDA 646

Qy     488 LLSKSVQLN-GLTLKMVDDQTL---PPLMEKPL 516
             | ||  |  :|| :: |  |   |||  : |
Db     647 APSDSVLSNVEVTLNVLADSVLMEQPPLRRRKL 679


RESULT 11
DYHC_RAT
ID   DYHC_RAT       STANDARD;      PRT;  4644 AA.
AC   P38650; Q63178;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Dynein heavy chain, cytosolic (DYHC) (Cytoplasmic dynein heavy chain)
DE   (MAP 1C).
GN   DNCHC1 OR DNCH1 OR DNEC1 OR MAP1C.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=Wistar; TISSUE=Brain;
RX   MEDLINE=93376715; PubMed=7690137;
RA   Zhang Z., Tanaka Y., Nonaka S., Aizawa H., Kawasaki H., Nakata T.,
RA   Hirokawa N.;
RT   "The primary structure of rat brain (cytoplasmic) dynein heavy chain,
RT   a cytoplasmic motor enzyme.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:7928-7932(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=93264075; PubMed=7684232;
RA   Mikami A., Paschal B.M., Mazumdar M., Vallee R.B.;
RT   "Molecular cloning of the retrograde transport motor cytoplasmic
RT   dynein (MAP 1C).";
RL   Neuron 10:787-796(1993).
CC   -!- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
CC       MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
CC       ORGANELLES ALONG MICROTUBULES.
CC   -!- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
CC       INTERMEDIATE AND LIGHT CHAINS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D13896; BAA02996.1; -.
DR   EMBL; L08505; AAA41103.1; -.
DR   PIR; A38905; A38905.
DR   InterPro; IPR004273; Dynein_heavy.
DR   Pfam; PF03028; Dynein_heavy; 1.
KW   Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT   DOMAIN       48     69       COILED COIL (POTENTIAL).
FT   DOMAIN      179    200       COILED COIL (POTENTIAL).
FT   DOMAIN      453    476       COILED COIL (POTENTIAL).
FT   DOMAIN      541    564       COILED COIL (POTENTIAL).
FT   DOMAIN     1169   1201       COILED COIL (POTENTIAL).
FT   DOMAIN     1229   1250       COILED COIL (POTENTIAL).
FT   DOMAIN     1355   1371       COILED COIL (POTENTIAL).
FT   DOMAIN     2012   2040       MICROTUBULE-BINDING (POTENTIAL).
FT   DOMAIN     3187   3273       COILED COIL (POTENTIAL).
FT   DOMAIN     3394   3498       COILED COIL (POTENTIAL).
FT   DOMAIN     3735   3798       COILED COIL (POTENTIAL).
FT   NP_BIND    1904   1911       ATP (POTENTIAL).
FT   NP_BIND    2222   2229       ATP (POTENTIAL).
FT   NP_BIND    2593   2600       ATP (POTENTIAL).
FT   NP_BIND    2935   2942       ATP (POTENTIAL).
FT   CONFLICT   1024   1025       SR -> MP (IN REF. 2).
FT   CONFLICT   1772   1772       N -> D (IN REF. 2).
FT   CONFLICT   2098   2098       P -> A (IN REF. 2).
FT   CONFLICT   2139   2139       F -> V (IN REF. 2).
FT   CONFLICT   2175   2175       D -> A (IN REF. 2).
FT   CONFLICT   2185   2185       K -> Q (IN REF. 2).
FT   CONFLICT   2366   2366       L -> V (IN REF. 2).
FT   CONFLICT   2382   2382       T -> S (IN REF. 2).
FT   CONFLICT   2463   2463       G -> A (IN REF. 2).
FT   CONFLICT   3219   3219       A -> D (IN REF. 2).
FT   CONFLICT   4131   4131       R -> K (IN REF. 2).
FT   CONFLICT   4366   4366       F -> S (IN REF. 2).
FT   CONFLICT   4511   4511       A -> G (IN REF. 2).
SQ   SEQUENCE   4644 AA;  532240 MW;  8C6ABDBEDF875D82 CRC64;

  Query Match             3.5%;  Score 99;  DB 1;  Length 4644;
  Best Local Similarity   19.5%;  Pred. No. 1.6e+02;
  Matches  109;  Conservative   85;  Mismatches  188;  Indels  178;  Gaps   25;

Qy      41 LDFFTQEPLHLVSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTD 100
           :|   ::|   | | |   :| |     |       :|| |       | | | | :

Db    1341 IDQMKEQPWVSVQPRKLRQNLDGLLNQLKNF-----PARLRQYASYEFVQRLLKGYMKIN 1395

Qy     101 FLIFDPKKESTFEERSYWQSQVNQ-----------------DICKYGSIPPDV------ 136
           |: : | |:  ::| :|:  : :                  |: |  :|  ||
Db    1396 MLVIELKSEA-LKDR-HWKQLMKRLHVNWVVSELTLGQIWDVDLQKNEAIVKDVLLVAQG 1453

Qy     137 ----EE---KLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLN 189
               ||   ::|  |   |  |:  :|| | :                    || | :
Db    1454 EMALEEFLKQIREVWNTYELDLV--NYQNKCR---------------------LIRGWD 1489

Qy     190 ALLRTADLQWNSSNAQLLLDYCSSKGYNI------SWELGNEPNSFLKKADIFIN----- 238
           |      || :|  |  |      | :      ||| :: | :   |::|:
Db    1490 DLFNKVKEHINSVSAMKLSPY-----YKVFEEDALSWE--DKLNRIMALFDVWIDVQRRW 1542

Qy     239 --------GS--------------QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKT 275
                  ||                :  ::: | | : ||      |    | :   :
Db    1543 VYLEGIFTGSADIKHLLPVETQRFQSISTEFLALMKKVSKSPLVMDVLNIQGVQRSLERL 1602

Qy     276 AKMLKSFLKAGGEVI--DSVTWHHYYLNGRTATREDFLNPDVLDIF-----ISSVQKVFQ 328
           | :|    || || :  :  ::  :|  |         : |:|:|      :: :|| |
Db    1603 ADLLGKIQKALGEYLERERSSFPRFYFVG---------DEDLLEIIGNSKNVAKLQKHF- 1652

Qy     329 VVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGA 388
                  ||:: | :|      : ::             ||:|:| | ||: :
Db    1653 --------KKMFAGVSSIILNEDSSVV------------LGISSREGEEVMFKTP---- 1687

Qy     389 GNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRY--- 445
                |     |  : ||:|: |::   :| :| :      :: ::   |: |   |
Db    1688 -----VSITEHPK1NEWLTLVEKEM---RVTLAKLLAESVTEVEIFGKATSIDPNTYITW 1739

Qy     446 ---KEGDLTLYAINL---HNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLT 499
               :  | : :  :    ||   |       |          :||   :  :|:   :|
Db    1740 IDKYQAQLVVLSAQIAWSENVENALSNVGGGGN----------VGPLQSVLSNVE---VT 1786

Qy     500 LKMVDDQTL---PPLMEKPL 516
           | :: |  |   |||  : |
Db    1787 LNVLADSVLMEQPPLRRRKL 1806


RESULT 12
T3MH_HAEIN
ID   T3MH_HAEIN     STANDARD;      PRT;   629 AA.
AC   P71366;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative type III restriction-modification system HindVIP enzyme mod
DE   (EC 2.1.1.72) (HindVIP methyltransferase) (M.HindVIP).
GN   HI1056.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC       adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC   -!- SUBUNIT: CONTAINS TWO DIFFERENT SUBUNITS: RES AND MOD. MOD IS
CC       A HOMOTETRAMER (BY SIMILARITY).
```

```
CC   -!- SIMILARITY: WITH OTHER TYPE III MOD PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32786; AAC22721.1; -.
DR   REBASE; 3701; M.HindORF1056P.
DR   TIGR; HI1056; -.
DR   InterPro; IPR001091; CN4_Metransf.
DR   InterPro; IPR002295; D21N6_mtfrase.
DR   InterPro; IPR002941; N6/N4_Mtase.
DR   InterPro; IPR002052; N6_Mtase.
DR   Pfam; PF01555; N6_N4_Mtase; 1.
DR   PRINTS; PR00506; D21N6MTFRASE.
DR   PRINTS; PR00508; S21N4MTFRASE.
DR   PROSITE; PS00092; N6_MTASE; 1.
KW   Hypothetical protein; Transferase; Methyltransferase; DNA-binding;
KW   Restriction system; Complete proteome.
SQ   SEQUENCE   629 AA;  71845 MW;  93ADAD909DB41E84 CRC64;

  Query Match             3.5%;  Score 98.5;  DB 1;  Length 629;
  Best Local Similarity   21.8%;  Pred. No. 9.8;
  Matches   67;  Conservative   56;  Mismatches  107;  Indels   77;  Gaps   19;

Qy      99 TDFL-IFDPKKESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQ-------EQL 150
           |:|: :|   :|  |    | :::||: |  |         |     | |         |:: 
Db     242 TEFINVFSKNRERLFLNPVYQKTEVNEFIKNY--------EDSGKSWKYTQVLIDLGEKI 293

Qy     151 LLREHYQKKFKNSTYSRSSVDVLYTFANCSGL--DLIFG--LNALLRTADLQ-------- 198
           || |  :  ||   |  : :  :  |:    |  ::|:    : : || : |
Db     294 LLEE--KDGFKYYHYPNAQMTSIVKFSQDQNLSKEIIYTEYSHKVYRTTNAQSSIRSKII 351

Qy     199 ---WNSSNAQLLLDYCSSKGYN------ISWELGNEPNSFLKKADIFING---------- 239
             ::  |  : ::|   || |      : :   |: : |:  :|: |
Db     352 EDLYSIKNGIVSIEYIPQKGKNAGNLIEVFYNASNK-DMFMFLSDMLIKEKNKYFYLQKV 410

Qy     240 SQLGEDYIQLHKLLRKS---TFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWH 296
           : | :| || : | ::     ||| |       :|     :::    | |  |:|
Db     411 NTLWDD-IQYNNLNKEGGYIDFKNGK-------KPEALLRRIIDMTTKEGDIVLD----- 457

Qy     297 HYYL-NGRTATREDFLNPDVLDI----FIS--SVQKVFQVVESTRPG--KKV-WLGETSS 346
            |:| :| ||     :|   : |    :|   :|::: :|::  : |  | | | |
Db     458 -YHLGSGTTAAVAHKMNRQYIGIEQMDYIETLAVERLKKVIDGEQGGISKAVNWQGGGEF 516

Qy     347 AYGGGAP 353
             |   ||
Db     517 VYAELAP 523


RESULT 13
NU5M_RHIST
ID   NU5M_RHIST     STANDARD;      PRT;   654 AA.
AC   P50367;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN   ND5 OR NAD5.
OS   Rhizopus stolonifer (Rhizopus nigricans).
OG   Mitochondrion.
OC   Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC   Rhizopus.
OX   NCBI_TaxID=4846;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DAOM 148428;
RA   Paquin B., Roewer I., Wang Z., Lang B.F.;
RT   "A robust fungal phylogeny using the mitochondrially encoded nad5
RT   protein sequence.";
RL   Can. J. Bot. 73:S180-S185(1995).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U17011; AAA99062.1; -.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ   SEQUENCE   654 AA;  72818 MW;  22D2E025B8D6E12D CRC64;

  Query Match             3.5%;  Score 98.5;  DB 1;  Length 654;
  Best Local Similarity   21.3%;  Pred. No. 10;
  Matches   77;  Conservative   55;  Mismatches  132;  Indels   97;  Gaps   17;

Qy     176 FANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADI 235
           :: || | |:| :                        |    ||::  | :  |    || :
Db     304 YSTCSQLGLLFLV--------------------CGLSQYNVA--LFHLVNHAWFKALL 339

Qy     236 FINGSQL------GEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEV 289
           |::   :       :|  :   | |   |  : :    :|         |  |     ::
Db     340 FLSAGSVIHAMNDEQDLRKFGGLSRLLPFTYSMMV---IGSLSLMALPFLTGFYSK--DL 394

Qy     290 IDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYG 349
           |  : : ||  :|              |  :::||  ||  : | |     :||     |
Db     395 IIELAYGHYSFSGN-----------LVYWLASVAAVFTAMYSIRSLYLTFLG-----YP 437

Qy     350 GG----------APLLS-------DTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYH 392
           |          |||:          |:  | :: |  |   || :     :|    :
Db     438 NGPKINYNNIHEAPLIMAIPLVVLAVFSIFFGYVTK-DLFVGMGTDFYNNALFIHPNHSI 496

Qy     393 LVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVY---------LHCTNTDNP 443
           |||  |  ||   :|: |  |:|: :    |       | :|           |
Db     497 LVDTEFG-LP---MSMKFLPLIGSLLGTFGV-------LAIYWIFDELPNKFISTKLGRG 545

Qy     444 RYKEGDLTLYAINLHN--VTKYLRLPYPFSNKQVDKYLLRPLGPHGLL----SKSVQLNG 497
           |:  :   |  |::|  : |:|   |  :|| :|:  :  :||:||:    | | :::|
Db     546 IYRFFNQKYYFDNIYNNLLNKFLNFGYT-TNKILDRGAIELVGPYGLVNVFKSASNKVSG 604

Qy     498 L 498
           |
Db     605 L 605


RESULT 14
RHTA_RHIME
ID   RHTA_RHIME     STANDARD;      PRT;   746 AA.
AC   Q9Z3Q5;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Rhizobactin receptor precursor (TonB-dependent siderophore receptor
DE   rhtA).
GN   RHTA OR RA1265 OR SMA2414.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymA (megaplasmid 1).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RCR2011 / SU47;
```

```
RX   MEDLINE=21172875; PubMed=11274118;
RA   Lynch D., O'Brien J., Welch T., Clarke P., Cuiv P.O., Crosa J.H.,
RA   O'Connell M.;
RT   "Genetic organization of the region encoding regulation, biosynthesis,
RT   and transport of rhizobactin 1021, a siderophore produced by
RT   Sinorhizobium meliloti.";
RL   J. Bacteriol. 183:2576-2585(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396509; PubMed=11481432;
RA   Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA   Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA   Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA   Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA   Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT   "Nucleotide sequence and predicted functions of the entire
RT   Sinorhizobium meliloti pSymA megaplasmid.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC   -!- FUNCTION: RECEPTOR FOR THE SIDEROPHORE RHIZOBACTIN.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF110737; AAD09419.1; -.
DR   EMBL; AE007312; AAK65923.1; -.
DR   InterPro; IPR000531; TonB_boxC.
DR   Pfam; PF00593; TonB_boxC; 1.
DR   PROSITE; PS00430; TONB_DEPENDENT_REC_1; FALSE_NEG.
DR   PROSITE; PS01156; TONB_DEPENDENT_REC_2; FALSE_NEG.
KW   Outer membrane; Iron transport; Transport; TonB box; Signal; Receptor;
KW   Plasmid; Complete proteome.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    746       RHIZOBACTIN RECEPTOR.
FT   SITE        40     47       TONB BOX.
FT   SITE       729    746       TONB C-TERMINAL BOX.
SQ   SEQUENCE   746 AA;  80633 MW;  16AE44A4025D5B23 CRC64;

  Query Match             3.5%;  Score 98.5;  DB 1;  Length 746;
  Best Local Similarity   21.8%;  Pred. No. 13;
  Matches  105;  Conservative   55;  Mismatches  159;  Indels  163;  Gaps   25;

Qy      90 AYLRFGGTKT--------DFLIFDPKKEST-FEERSYWQSQVNQDICKYGSIPPDVEEKL 140
           | |   | :|          || |  : || | ||          |   |||   :::
Db     207 ARLSIAGNRTGAFYDGSGTLLIPDITQTSTAFNER----------IDLMGSIGYQIDDDR 256

Qy     141 RLEW-----------------PYQEQLL------LREHYQKKFKNSTYSRSSVDVLYTF 176
           |:|:                  |:  |        |  |:  |   |  || ::| ||
Db     257 RVEFSGQYFDSKQDSDYGLYYGPFFAALADPSLFETRSGYESDFNPQT-RRSMLNVTYTD 315

Qy     177 ANCSGLDLIFGLNALLRTADLQWN----SSNAQL-LLDYCSSKG---YNISWELGNEP-- 226
             :  |  |:  |    ||  ::::    | |::     | ||:    | |   |  ||
Db     316 NDVFGQQLL--LQGSYRTERIKFHPFPASGNSETGPYFYGSSQDTDYYGIRAALVAEPTD 373

Qy     227 ------------NSFLKKADIF---INGSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQP 271
                       :||  : :||      |   | |:         :|     || | :
Db     374 ALKITYGIDADMDSFTARQNIFDMVAAGQSGGLDF---------NTIGKTGLY-PSI--- 420

Qy     272 RRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVE 331
           :  | :|  |  | :|     |||    |  |:| :| | || : |:|
Db     421 ---DVSTVAGFAEASYEATDRLT-----LNG--GVRYQFVNTEVSD-FIGAAQQV----- 464

Qy     332 STRPGKKVWLGETSSA---YGGGAPLLSDTFAAGFMW----LDKLGLSARMGIEVVMRQV 384
           :  | :||    ||      :  |:||  :      ::  :   | |:
Db     465 ------AILQGRATSADTIPGGEVNYDAALFSAGATYQLTNTQQVYANFSQGFELPDPAK 518

Qy     385 FFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPR 444
           ::| |||                          |   |: ||
Db     519 YYGIGNYSFSG----------------GHYTLVNSV---------------------- 538

Qy     445 YKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLN--GLTLKM 502
             ||  | ||  ::     ||     | :  |       : |  :|: ||   | :::
Db     539 -NVGDSALEAIKTNSFEIGYRLDDGTFNLETAAY-------YSLSDRSINLNRSSLAVEI 590

Qy     503 VD 504
           :|
Db     591 ID 592


RESULT 15
RRP2_IAVI7
ID   RRP2_IAVI7     STANDARD;      PRT;   716 AA.
AC   P31343;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   RNA-directed RNA polymerase subunit P2 (EC 2.7.7.48) (Polymerase
DE   acidic protein) (PA).
OS   Influenza A virus (strain A/Victoria/3/75).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=11483;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89370813; PubMed=2773594;
RA   de la Luna S., Martinez C., Ortin J.;
RT   "Molecular cloning and sequencing of influenza virus A/Victoria/3/75
RT   polymerase genes: sequence evolution and prediction of possible
RT   functional domains.";
RL   Virus Res. 13:143-156(1989).
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: INFLUENZA RNA POLYMERASE IS COMPOSED OF THREE SUBUNITS:
CC       P1 (OR PB1), P2 (OR PA), AND P3 (OR PB2).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA VIRUSES POLYMERASE PA FAMILY.
DR   PIR; C60008; C60008.
DR   InterPro; IPR001009; RNA_pol_P2.
DR   Pfam; PF00603; Flu_PA; 1.
KW   Transferase; RNA-directed RNA polymerase.
SQ   SEQUENCE   716 AA;  82821 MW;  B1FC817C171BB50F CRC64;

  Query Match             3.4%;  Score 98;  DB 1;  Length 716;
  Best Local Similarity   19.3%;  Pred. No. 13;
  Matches  100;  Conservative   80;  Mismatches  194;  Indels  144;  Gaps   25;

Qy      57 LSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERS 116
           ::  |:  | | || : |   |           | : |     : ||: |  | |  :
Db     253 VNAKIEPFLKTTPRPIKLPDGP----------PCFQR-----SKFLLMDALKLSIED--- 294

Qy     117 YWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTF 176
             |   : |  | :|     : :|  : ::|  ::: |  ::  || |  |   ||
Db     295 --PSHEGEGIPLYDAI-----KCMRTFFGWKEPYIVKPH--ERGINSNYLLSWKQVLAEL 345

Qy     177 ANCSGLDLIFGLNALLRTADLQW----NSSNAQLLLDYCSS----KGYN---------IS 219
             :    : |     : :|: |:|    | :  ::  | |       | |:         |
Db     346 QDIENEEKIPRTKNMKKTSQLKWALGENMAPEKVDFDNCRDISDLKQYDSDEPELRSLSS 405

Qy     220 WELGNEPNSFLKKAD-IFINGSQLGEDYIQLHKL--LRKSTFKNAKLYGPDVGQPRRKTA 276
           | : || |   :  | |:|   ::|||   :  :  :|:: |                 ||
Db     406 W-IQNEFNKACELTDSIWIELDEIGEDVAPIEYIASMRRNYF---------------TA 448

Qy     277 KMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPG 336
           ::  |  :|   ::  |  :   ||   |  :||         | : |       |: |
Db     449 EV--SHCRATEYIMKGVYINTALLNASCAAMDDFQ-------LIPMISKC-----RTKEG 494

Qy     337 KKVWLGETSSAYG----GGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYH 392
           ::      :: ||    | : | :||    |: ::                      :
Db     495 RR-----KTNLYGFIIKGRSHLRNDTDVVNFVSME----------------------FS 526
```

```
Qy     393 LVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTL 452
           | |   :  |  |      ::    :|: |  |   | : :|:    |   : | |
Db     527 LTDPRLE--PHKWEKYCVLEI--GDMLLRSAIGQMSRPMFLYVRTNGTSKIKMKWGMEMR 582

Qy     453 YAI--NLHNVTKYLRLPYPFSNKQVDKYLLR------PLG--PHG-----------LLS 490
             :  :|  :   :        | : |          |:|  | |            ||:
Db     583 RCLLQSLQQIESMIEAESSVKEKDMTKEFFENKSETWPIGESPKGVEEGSIGKVCRTLLA 642

Qy     491 KSV--------QLNGLTLKMVDDQTLPPLMEKPLRPGS 520
           |||        || | : :      :   :   | ||:
Db     643 KSVFNSLYASPQLEGFSAESRKLLLVVQALRDNLEPGT 680
```

Search completed: November 20, 2002, 11:36:35
Job time : 19 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 20, 2002, 11:34:52 ; Search time 39 Seconds
(without alignments)
2868.810 Million cell updates/sec

Title: US-09-759-207-2
Perfect score: 2842
Sequence: 1 MLLRSKPALPPPLMLLLLGP.....LPAFSYSFFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2838	99.9	543	4	Q9Y251	Q9y251 homo sapien
2	2817	99.1	545	4	Q9UL39	Q9ul39 homo sapien
3	2282	80.3	545	6	Q9MYY0	Q9myy0 bos taurus
4	2123	74.7	536	11	Q9QZF8	Q9qzf8 rattus norv
5	1645.5	57.9	523	13	Q90YK5	Q90yk5 gallus gall
6	1154.5	40.6	592	4	Q9HB37	Q9hb37 homo sapien
7	1146.5	40.3	592	4	Q8WWQ2	Q8wwq2 homo sapien
8	1015.5	35.7	548	4	Q8WWQ1	Q8wwq1 homo sapien
9	936.5	33.0	534	4	Q9HB38	Q9hb38 homo sapien
10	897.5	31.6	480	4	Q9HB39	Q9hb39 homo sapien
11	696	24.5	515	5	Q8T108	Q8t108 bombyx mori
12	416	14.6	521	10	Q9SDA1	Q9sda1 arabidopsis
13	416	14.6	543	10	Q9FF10	Q9ff10 arabidopsis
14	381	13.4	527	10	Q9LRC8	Q9lrc8 scutellaria
15	363	12.8	536	10	Q9FZP1	Q9fzp1 arabidopsis
16	352.5	12.4	516	10	Q9FLK8	Q9flk8 arabidopsis
17	169.5	6.0	190	10	O82604	O82604 arabidopsis
18	160	5.6	935	5	Q9VE79	Q9ve79 drosophila
19	130.5	4.6	493	17	Q9HK01	Q9hk01 thermoplasm
20	122.5	4.3	408	3	Q9HEZ2	Q9hez2 phanerochae
21	122.5	4.3	408	3	Q9HEZ1	Q9hez1 phanerochae
22	116.5	4.1	617	12	O40996	O40996 measles vir
23	113.5	4.0	617	12	Q83295	Q83295 measles vir
24	113.5	4.0	1829	2	Q9KH44	Q9kh44 pantoea agg
25	113	4.0	390	17	Q8TPH7	Q8tph7 methanosarc
26	112.5	4.0	617	12	Q83647	Q83647 measles vir
27	111.5	3.9	493	17	Q979W0	Q979w0 thermoplasm
28	111.5	3.9	575	10	Q43855	Q43855 vicia faba
29	111.5	3.9	617	12	O40991	O40991 measles vir
30	111.5	3.9	617	12	Q98VH5	Q98vh5 measles vir
31	111	3.9	670	10	Q9M090	Q9m090 arabidopsis
32	111	3.9	2319	3	Q96U00	Q96u00 neurospora
33	110.5	3.9	475	5	Q8SU17	Q8su17 encephalito
34	109.5	3.9	617	12	Q91IP6	Q91ip6 measles vir
35	109.5	3.9	617	12	Q910N9	Q910n9 measles vir
36	109.5	3.9	879	16	Q8XCP4	Q8xcp4 escherichia
37	109	3.8	617	12	Q98VT6	Q98vt6 measles vir
38	108.5	3.8	411	16	P72895	P72895 synechocyst
39	108.5	3.8	617	12	O11381	O11381 measles vir
40	107.5	3.8	500	16	Q9A5U0	Q9a5u0 caulobacter
41	107.5	3.8	617	12	Q89631	Q89631 measles vir
42	107.5	3.8	617	12	Q89764	Q89764 measles vir
43	107.5	3.8	617	12	O40990	O40990 measles vir
44	107.5	3.8	617	12	Q83633	Q83633 measles vir
45	107.5	3.8	617	12	Q83645	Q83645 measles vir

ALIGNMENTS

RESULT 1
Q9Y251
ID Q9Y251 PRELIMINARY; PRT; 543 AA.
AC Q9Y251;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE HEPARANASE.
GN HPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=99321249; PubMed=10395326;
RA Hulett M.D., Freeman C., Hamdorf B.J., Baker R.T., Harris M.J.,
RA Parish C.R.;
RT "Cloning of mammalian heparanase, an important enzyme in tumor
RT invasion and metastasis.";
RL Nat. Med. 5:803-809(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Vlodavsky I., Friedman Y., Elkin M., Aingorn H., Atzmon R.,
RA Ishai-Michaeli R., Bitan M., Pappo O., Peretz T., Michal I.,
RA Spector L., Pecker I.;
RT "Mammalian heparanase: a novel gene involved in tumor progression and
RT metastasis.";
RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99377052; PubMed=10446189;
RA Toyoshima M., Nakajima M.;
RT "Human heparanase. Purification, characterization, cloning, and
RT expression.";
RL J. Biol. Chem. 274:24153-24160(1999).
RN [4]
RP SEQUENCE FROM N.A.

```
RC   TISSUE=PLACENTA;
RX   MEDLINE=99335379; PubMed=10405343;
RA   Kussie P.H., Hulmes J.D., Ludwig D., Patel S., Navarro E.C.,
RA   Seddon A.P., Giorgio N.A., Bohlen P.;
RT   "Cloning and Functional Expression of a Human Heparanase Gene.";
RL   Biochem. Biophys. Res. Commun. 261:183-187(1999).
DR   EMBL; AF165154; AAD45379.1; -.
DR   EMBL; AF144325; AAD41342.1; -.
DR   EMBL; AF155510; AAD54941.1; -.
DR   EMBL; AF152376; AAD45669.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   543 AA;  61176 MW;  AD262EC267334AB2 CRC64;

  Query Match             99.9%;  Score 2838;  DB 4;  Length 543;
  Best Local Similarity   99.8%;  Pred. No. 7.8e-214;
  Matches  542;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy     541 ACI 543
           |||
Db     541 ACI 543


RESULT 2
Q9UL39
ID   Q9UL39      PRELIMINARY;      PRT;   545 AA.
AC   Q9UL39;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=20229546; PubMed=10764835;
RA   Dempsey L.A., Plummer T.B., Coombes S.L., Platt J.L.;
RT   "Heparanase expression in invasive trophoblasts and acute vascular
RT   damage.";
RL   Glycobiology 10:467-475(2000).
DR   EMBL; AF084467; AAD54516.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   545 AA;  61418 MW;  67B80ACD73C5A9A1 CRC64;

  Query Match             99.1%;  Score 2817;  DB 4;  Length 545;
  Best Local Similarity   99.4%;  Pred. No. 3.5e-212;
  Matches  542;  Conservative    1;  Mismatches    0;  Indels    2;  Gaps    2;

Qy       1 MLLRSKPALPPP-LMLLLLGPLGPLSPGALPRPAQA-QDVVDLDFFTQEPLHLVSPSFLS 58
           |||||||||||| |||||||||||||||||||||||| |||||||||||||||||||||||
Db       1 MLLRSKPALPPPLLMLLLLGPLGPLSPGALPRPAQAQQDVVDLDFFTQEPLHLVSPSFLS 60

Qy      59 VTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYW 118
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYW 120

Qy     119 QSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFAN 178
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFAN 180

Qy     179 CSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFIN 238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFIN 240

Qy     239 GSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHY 298
           |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GSQLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHY 300

Qy     299 YLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDT 358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 YLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDT 360

Qy     359 FAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKV 418
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 FAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKV 420

Qy     419 LMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKY 478
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 LMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKY 480

Qy     479 LLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAK 538
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 LLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAK 540

Qy     539 VAACI 543
           |||||
Db     541 VAACI 545


RESULT 3
Q9MYY0
ID   Q9MYY0      PRELIMINARY;      PRT;   545 AA.
AC   Q9MYY0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Kizaki K., Nakano H., Takahashi T., Imai K., Hashizume K.;
```

```
RT   "Expression of Heparanase mRNA in Bovine Placenta During Gestation.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF281160; AAF87301.2; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   545 AA;  61076 MW;  FAC4BDFFD855B933 CRC64;

  Query Match             80.3%;  Score 2282;  DB 6;  Length 545;
  Best Local Similarity   79.8%;  Pred. No. 2.9e-170;
  Matches  435;  Conservative   35;  Mismatches   73;  Indels    2;  Gaps    1;

Qy       1 MLLRSKPALPPPLMLL--LLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLS 58
           ||   || | |||:||  ||||||| |||     | | |  :|:|||: |||||||:|||
Db       1 MLACRKPGLRPPLLLLLPLLGPLGPCSPGTPAAAAPADDAAELEFFTERPLHLVSPAFLS 60

Qy      59 VTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYW 118
           ||||||||||||   ||| ||||||||:||||||| | ||||||||||  |||||||
Db      61 FTIDANLATDPRFFTFLGSSKLRTLARGLAPAYLRFGGNKGDFLIFDPKKEPAFEERSYW 120

Qy     119 QSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFAN 178
           || |||||| |||| ||||||||||:|||:|||| ||||| ||||||||||:|||||:
Db     121 LSQSNQDICKSGSIPSDVEEKLRLEWPFQEQVLLREQYQKKFTNSTYSRSSVDMLYTFAS 180

Qy     179 CSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFIN 238
           ||||:||||:|||||| |: |:|||||||||||| |||||||||||| :|| ||||
Db     181 CSGLNLIFGVNALLRTTDMHWDSSNAQLLLDYCSSKNYNISWELGNEPNSFQRKAGIFIN 240

Qy     239 GSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHY 298
           | ||||||:|:  ||| || ||||||||||:||||| | ||||||||||||||||||||
Db     241 GRQLGEDFIEFRKLLGKSAFKNAKLYGPDIGQPRRNTVKMLKSFLKAGGEVIDSVTWHHY 300

Qy     299 YLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDT 358
           |:||| ||:||||||:|| |||||||  ::||  || ||||||||||:||||| ||:|
Db     301 YVNGRIATKEDFLNPDILDTFISSVQKTLRIVEKIRPLKKVWLGETSSAFGGGAPFLSNT 360

Qy     359 FAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKV 418
           |||||||||||||||||||||||| |||||||||| ||:|||||||||||||||| ||
Db     361 FAAGFMWLDKLGLSARMGIEVVMRQVLFGAGNYHLVDGNFEPLPDYWLSLLFKKLVGNKV 420

Qy     419 LMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKY 478
           |||||:|  | | |||||||||| :|||||||||||:|||||||:| ||:   |||||||
Db     421 LMASVKGPDRSKFRVYLHCTNTKHPRYKEGDLTLYALNLHNVTKHLELPHHLFNKQVDKY 480

Qy     479 LLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAK 538
           |::| |  |||||||||||  |||||:|||| | |||| |||||:| ||| ||||||||
Db     481 LIKPSGTDGLLSKSVQLNGQILKMVDEQTLPALTEKPLHPGSSLGMPPFSYGFFVIRNAK 540

Qy     539 VAACI 543
           |||||
Db     541 VAACI 545


RESULT 4
Q9QZF8
ID   Q9QZF8      PRELIMINARY;      PRT;   536 AA.
AC   Q9QZF8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase.
GN   HEP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Podyma K.A., Yokote H., Sakaguchi K., Ikuta M., Yanagishita M.;
RT   "Heparanase from parathyroid cell line.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF184967; AAF04563.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   536 AA;  60569 MW;  6208B1FD9EE28421 CRC64;

  Query Match             74.7%;  Score 2123;  DB 11;  Length 536;
  Best Local Similarity   75.7%;  Pred. No. 8e-158;
  Matches  405;  Conservative   51;  Mismatches   79;  Indels    0;  Gaps    0;

Qy       9 LPPPLMLLLLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLSVTIDANLATD 68
           | | |:| | | |  |: |     |  :|||||:|:|:     |||||||:||||:||||
Db       2 LRPLLLLWLWGRLRALTQGTPAGTAPTKDVVDLEFYTKRLFQSVSPSFLSITIDASLATD 61

Qy      69 PRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICK 128
           ||||  ||||:|| |||||||||||||||||||||||| || | |||||||| | |||
Db      62 PRFLTFLGSPRLRALARGLSPAYLRFGGTKTDFLIFDPNKEPTSEERSYWQSQDNNDICG 121

Qy     129 YGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGL 188
             :  ||  ||::|||:|| ||||| ||::|||||||||:||:|| || |||||||
Db     122 SERVSADVLRKLQMEWPFQELLLLREQYQREFKNSTYSRSSVDMLYSFAKCSRLDLIFGL 181

Qy     189 NALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDYIQ 248
           |||||| ||:|||||||||:|||||||||||||||||||| ||| | |:| |||||:::
Db     182 NALLRTPDLRWNSSNAQLLLNYCSSKGYNISWELGNEPNSFWKKAQISIDGLQLGEDFVE 241

Qy     249 LHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATRE 308
           |||||:|| |:|||||||:|||| || |:|:|||||||||||:|||||||||| ||:|
Db     242 LHKLLQKSAFQNAKLYGPDIGQPRGKTVKLLRSFLKAGGEVIDSLTWHHYYLNGRVATKE 301

Qy     309 DFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDK 368
           |||: |||| || ||||: :| :   |||||||||||||||||||||:||||||||||
Db     302 DFLSSDVLDTFILSVQKILKVTKEMTPGKKVWLGETSSAYGGGAPLLSNTFAAGFMWLDK 361

Qy     369 LGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKR 428
           |||||::|||||||||||||||||||:|||||||||||||||| |||| |:|  |
Db     362 LGLSAQLGIEVVMRQVFFGAGNYHLVDENFEPLPDYWLSLLFKKLVGPKVLMSRVKGPDR 421

Qy     429 RKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGL 488
           |||||||||| :|||:|||||| :|||||||:|:|| |  :: ||||||:| |  ||
Db     422 SKLRVYLHCTNVYHPRYREGDLTLYVLNLHNVTKHLKLPPPMFSRPVDKYLLKPFGSDGL 481

Qy     489 LSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI 543
           ||||||||| |||||:|||| | ||||  |||| :||||| ||||||||:||||
Db     482 LSKSVQLNGQTLKMVDEQTLPALTEKPLPAGSSLSVPAFSYGFFVIRNAKIAACI 536


RESULT 5
Q90YK5
ID   Q90YK5      PRELIMINARY;      PRT;   523 AA.
AC   Q90YK5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11387326;
RA   Goldshmidt O., Zcharia E., Aingorn H., Guatta-Rangini Z., Atzmon R.,
RA   Michal I., Pecker I., Mitrani E., Vlodavsky I.;
RT   "Expression Pattern and Secretion of Human and Chicken Heparanase Are
RT   Determined by Their Signal Peptide Sequence.";
RL   J. Biol. Chem. 276:29178-29187(2001).
DR   EMBL; AY037007; AAK82648.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   523 AA;  58386 MW;  8EB0B7B18C9BF881 CRC64;

  Query Match             57.9%;  Score 1645.5;  DB 13;  Length 523;
  Best Local Similarity   60.2%;  Pred. No. 2e-120;
  Matches  320;  Conservative   87;  Mismatches  114;  Indels   11;  Gaps    3;
```

```
Qy      13 LMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFL 72
           |:||||  |  : |        :   :|    :||:  |||:|||:|:||:|| ||||:
Db       2 LVLLLLVLLLAVPP---------RRTAELQLGLREPIGAVSPAFLSLTLDASLARDPRFV 52

Qy      73 ILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSI 132
           ||  ||| ||| |||| :|||||| |||||||:| |:||:||:   : |  :|:|:
Db      53 ALLRHPKLHTLASGLSPGFLRFGGTSTDFLIFNPNKDSTWEEKVLSEFQA-KDVCEAWPS 111

Qy     133 PPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALL 192
           | : |  :|| ||:||| ||  || ||:| :|||::|:|:|||: ||  |:|||||||
Db     112 FAVVPKLLLTQWPLQEKLLLAEHSWKKHKNTTITRSTLDILHTFASSSGFRLVFGLNALL 171

Qy     193 RTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDYIQLHKL 252
           | | |||:||||: || ||: : |||||||||||||| ||: | |:| ||| |:: | :|
Db     172 RRAGLQWDSSNAKQLLGYCAQRSYNISWELGNEPNSFRKKSGICIDGFQLGRDFVHLRQL 231

Qy     253 L-RKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFL 311
           | :   :::|:||| ||||||: |  :|:||:|:||: |||||||||:|||:|||||||
Db     232 LSQHPLYRHAELYGLDVGQPRKHTQHLLRSFMKSGGKAIDSVTWHHYYVNGRSATREDFL 291

Qy     312 NPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGL 371
           :|:||| | :::  |  :||:| |||||||||| |||||||| ||:|: ||||||||||
Db     292 SPEVLDSFATAIHDVLGIVEATVPGKKVWLGETGSAYGGGAPQLSNTYVAGFMWLDKLGL 351

Qy     372 SARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKL 431
           :|| ||:|||||| ||||:|||||  | ||||||||||:|:||||:|| ||| :  |:
Db     352 AARRGIDVVMRQVSFGAGSYHLVDAGFKPLPDYWLSLLYKRLVGTRVLQASVEQADARRP 411

Qy     432 RVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSK 491
           ||||||||  :|:|:|||:||:|:|| |||: |:||    :| ||:||| | |   :||:
Db     412 RVYLHCTNPRHPKYREGDVTLFALNLSNVTQSLQLPKQLWSKSVDQYLLLPHGKDSILSR 471

Qy     492 SVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI 543
           |||||  |:||||:||| | |  | |||:|||||||| |:|||||  |||
Db     472 EVQLNGRLLQMVDDETLPALHEMALAPGSTLGLPAFSYGFYVIRNAKAIACI 523


RESULT 6
Q9HB37
ID   Q9HB37      PRELIMINARY;      PRT;   592 AA.
AC   Q9HB37;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase-like protein HPA2c.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20483645; PubMed=11027606;
RA   McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA   Hircock M., Patel S., Barry E., Stubberfield C., Terrett J., Page M.;
RT   "Cloning and Expression Profiling of Hpa2, a Novel Mammalian
RT   Heparanase Family Member.";
RL   Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
DR   EMBL; AF282887; AAG23423.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   592 AA;  66580 MW;  95C384AD9A74258E CRC64;

  Query Match             40.6%;  Score 1154.5;  DB 4;  Length 592;
  Best Local Similarity   44.2%;  Pred. No. 7.1e-82;
  Matches  251;  Conservative   81;  Mismatches  199;  Indels   37;  Gaps    8;

Qy      10 PPPLMLLLLGPLGPLSPGALPRP--------AQAQDVVDLDFFTQEPLHLVSPSFLSVTI 61
           |  | | ||  |   |      ||          : : :: ||  |: |:  |: :|||: :
Db      24 PGALYLALLLHLSLSSQAGDRRPLPVDRAAGLKEKTLILLDVSTKNPVRTVNENFLSLQL 83

Qy      62 DANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIF----DPKKESTFEERSY 117
           | ::  |  :|  | | :| ||||||||||:||||| :|||| |    :| |        |
Db      84 DPSIIHD-GWLDFLSSKRLVTLARGLSPAFLRFGGKRTDFLQFQNLRNPAKSRGGPGPDY 142

Qy     118 WQSQVNQDI-----------CKYGSIPPDVEEKLRLEWPYQEQL-LLREHYQKKFKNST 164
           :      ||            ||     ||| :|: |   |  | ||:| :   : |
Db     143 YLKNYEDDIVRSDVALDKQKGCKIAQ-HPDVMLELQREKAAQMHLVLLKEQFSNTYSNLI 201

Qy     165 YSRSSVDVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGN 224
           :  |:| || ||:|||| ||| |||| |  :  ||||:|  || | :|| ||||||||
Db     202 LTARSLDKLYNFADCSGLHLIFALNALRRNPNNSWNSSSALSLLKYSASKKYNISWELGN 261

Qy     225 EPNSFLKKADIFINGSQLGEDYIQLHKLLRK-STFKNAKLYGPDVGQPRRKTAKMLKSFL 283
           |||::       :||||||:|||||  ||:    :  | ||||::|:||:    :| |:
Db     262 EPNNYRTMHGRAVNGSQLGKDYIQLKSLLQPIRIYSRASLYGPNIGRPRKNVIALLDGFM 321

Qy     284 KAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGE 343
           |  |  :|:||| | |::||      |||   :||     ::|: :|| :  ||||:||
Db     322 KVAGSTVDAVTWQHCYIDGRVVKVMDFLKTRLLDTLSDQIRKIQKVVNTYTPGKKIWLEG 381

Qy     344 TSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPD 403
           :   ||   |||::||||:||: ||: |  ||:||:|  ||  |  ||||:||:||||
Db     382 VVTTSAGGTNNLSDSYAAGFLWLNTLGMLANQGIDVVIRHSFFDHGYNHLVDQNFNPLPD 441

Qy     404 YWLSLLFKKLVGTKVLMASVQGSKRR---------KLRVYLHCTNTDNPRYKEGDLTLYA 454
           ||||||:|:|:| |||   | | :|:          |||:| ||||  |  |  | :||:
Db     442 YWLSLLYKRLIGPKVLAVHVAGLQRKPRPGRVIRDKLRIYAHCTNHHNHNYVRGSITLFI 501

Qy     455 INLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEK 514
           ||||  | ::|     :| | :|||:| |  || ||||||||  | |||| ||| |  :
Db     502 INLHRSRKKIKLAGTLRDKLVHQYLLQPYGQEGLKSKSVQLNGQPLVMVDDGTLPELKPR 561

Qy     515 PLRPGSSLGLPAFSYSFFVIRNAKVAAC 542
           ||| | :| :|  :  |||::|.   ||
Db     562 PLRAGRTLVIPPVTMGFFVVKNVNALAC 589


RESULT 7
Q8WWQ2
ID   Q8WWQ2      PRELIMINARY;      PRT;   592 AA.
AC   Q8WWQ2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RA   Pessegue Safontas B.J.O.P.S.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RA   Legoux P., Legoux R., O'Brien D., Salome M.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ299719; CAC82491.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   592 AA;  66520 MW;  9478841FEACD558B CRC64;

  Query Match             40.3%;  Score 1146.5;  DB 4;  Length 592;
  Best Local Similarity   44.0%;  Pred. No. 3e-81;
  Matches  250;  Conservative   81;  Mismatches  200;  Indels   37;  Gaps    8;

Qy      10 PPPLMLLLLGPLGPLSPGALPRP--------AQAQDVVDLDFFTQEPLHLVSPSFLSVTI 61
           |  | | ||  |   |      ||          : : :: ||  |: |:  |: :|||: :
Db      24 PGALYLALLLHLSLSSQAGDRRPLPVDRAAGLKEKTLILLDVSTKNPVRTVNENFLSLQL 83

Qy      62 DANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIF----DPKKESTFEERSY 117
           | ::  |  :|  | | :| ||||||||||:||||| :|||| |    :| |        |
Db      84 DPSIIHD-GWLDFLSSKRLVTLARGLSPAFLRFGGKRTDFLQFQNLRNPAKSRGGPGPDY 142
```

```
Qy    118 WQSQVNQDI------------CKYGSIPPDVEEKLRLEWPYQEQL-LLREHYQKKFKNST 164
          :      ||             ||     |||  :|: |   |  | ||:| :   : |
Db    143 YLKNYEDDIVRSDVALDKQKGCKIAQ-HPDVMLELQREKAAQMHLVLLKEQFSNTYSNLI 201

Qy    165 YSRSSVDVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGN 224
          :  |:| ||  |:|||| ||| |||| |  :  ||||:|  || | :|| |||||||||
Db    202 LTARSLDKLYNSADCSGLHLIFALNALRRNPNNSWNSSSALSLLKYSASKKYNISWELGN 261

Qy    225 EPNSFLKKADIFINGSQLGEDYIQLHKLLRK-STFKNAKLYGPDVGQPRRKTAKMLKSFL 283
          |||::       :||||||:|||||  ||:    :  | ||||::|:||:    :|  |:
Db    262 EPNNYRTMHGRAVNGSQLGKDYIQLKSLLQPIRIYSRASLYGPNIGRPRKNVIALLDGFM 321

Qy    284 KAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGE 343
          |  |  :|:||| | |::||     |||   :||     ::|: :|| :  ||||:||
Db    322 KVAGSTVDAVTWQHCYIDGRVVKVMDFLKTRLLDTLSDQIRKIQKVVNTYTPGKKIWLEG 381

Qy    344 TSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPD 403
            :   ||   |||::|||||:||: ||: |  ||:||:|  ||  |  ||||:||:|||
Db    382 VVTTSAGGTNNLSDSYAAGFLWLNTLGMLANQGIDVVIRHSFFDHGYNHLVDQNFNPLPD 441

Qy    404 YWLSLLFKKLVGTKVLMASVQGSKRR---------KLRVYLHCTNTDNPRYKEGDLTLYA 454
          ||||||:|:|:| |||   | | :|:          |||:| ||||  |  |  | :|||:
Db    442 YWLSLLYKRLIGPKVLAVHVAGLQRKPRPGRVIRDKLRIYAHCTNHHNHNYVRGSITLFI 501

Qy    455 INLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEK 514
          ||||   | ::|     :| | :|||:| |  || ||||||||  | |||| ||| |  :
Db    502 INLHRSRKKIKLAGTLRDKLVHQYLLQPYGQEGLKSKSVQLNGQPLVMVDDGTLPELKPR 561

Qy    515 PLRPGSSLGLPAFSYSFFVIRNAKVAAC 542
          ||| | :| :|  :  |||::|    ||
Db    562 PLRAGRTLVIPPVTMGFFVVKNVNALAC 589


RESULT 8
Q8WWQ1
ID   Q8WWQ1      PRELIMINARY;      PRT;   548 AA.
AC   Q8WWQ1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase 3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RA   Pessegue Safontas B.J.O.P.S.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RA   Legoux P., Legoux R., O'Brien D., Salome M.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ299720; CAC82492.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ   SEQUENCE   548 AA;  61771 MW;  B8986303FC73A60A CRC64;

  Query Match             35.7%;  Score 1015.5;  DB 4;  Length 548;
  Best Local Similarity   41.8%;  Pred. No. 4.9e-71;
  Matches  224;  Conservative   79;  Mismatches  180;  Indels   53;  Gaps    9;

Qy      8 ALPPPLMLLLLGPLGPLSPGAL-----------------PRPA------QAQDVVDLDF 43
          | |  ::     |   |:||||                   | |       : : :: ||
Db      6 AFPEAMLSSNSRPPACLAPGALYLALLLHLSLSSQAGDRRPLPVDRAAGLKEKTLILLDV 65

Qy     44 FTQEPLHLVSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLI 103
          |: |:  |: :|||: :| ::  |  :|  | | :| ||||||||:|||| :||||
Db     66 STKNPVRTVNENFLSLQLDPSIIHD-GWLDFLSSKRLVTLARGLSPAFLRFGGKRTDFLQ 124

Qy    104 F----DPKKESTFEERSYWQSQVNQDI------------CKYGSIPPDVEEKLRLEWPYQ 147
          |    :| |         |:      ||            ||     |||  :|: |   |
Db    125 FQNLRNPAKSRGGPGPDYYLKNYEDDIVRSDVALDKQKGCKIAQ-HPDVMLELQREKAAQ 183

Qy    148 EQL-LLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQL 206
            | ||:| :   : |   :  |:| ||  |:|||| ||| |||| |  :  ||||:|
Db    184 MHLVLLKEQFSNTYSNLILTARSLDKLYNSADCSGLHLIFALNALRRNPNNSWNSSSALS 243

Qy    207 LLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDYIQLHKLLRK-STFKNAKLYG 265
          || | :|| ||||||||||||::        :||||||:|||||  ||:    :  | |||
Db    244 LLKYSASKKYNISWELGNEPNNYRTMHGRAVNGSQLGKDYIQLKSLLQPIRIYSRASLYG 303

Qy    266 PDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQK 325
          |::|:||:    :|  |:|  |  :|:||| | |::||     |||   :||     ::|
Db    304 PNIGRPRKNVIALLDGFMKVAGSTVDAVTWQHCYIDGRVVKVMDFLKTRLLDTLSDQIRK 363

Qy    326 VFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVF 385
          : :|| :  ||||:||    :   ||   |||::||||:||: ||: |  ||:||:|  |
Db    364 IQKVVNTYTPGKKIWLEGVVTTSAGGTNNLSDSYAAGFLWLNTLGMLANQGIDVVIRHSF 423

Qy    386 FGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRR---------KLRVYLH 436
          |  |  ||||:||:||||||||||:|:|:| |||   | | :|:          |||:| |
Db    424 FDHGYNHLVDQNFNPLPDYWLSLLYKRLIGPKVLAVHVAGLQRKPRPGRVIRDKLRIYAH 483

Qy    437 CTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKS 492
          |||  |  |  | :||: ||||   | ::|     :| | :|||:| |  || ||:
Db    484 CTNHHNHNYVRGSITLFIINLHRSRKKIKLAGTLRDKLVHQYLLQPYGQEGLKSKT 539


RESULT 9
Q9HB38
ID   Q9HB38      PRELIMINARY;      PRT;   534 AA.
AC   Q9HB38;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Heparanase-like protein HPA2b.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20483645; PubMed=11027606;
RA   McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA   Hircock M., Patel S., Barry E., Stubberfield C., Terrett J., Page M.;
RT   "Cloning and Expression Profiling of Hpa2, a Novel Mammalian
RT   Heparanase Family Member.";
RL   Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
DR   EMBL; AF282886; AAG23422.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
SQ   SEQUENCE   534 AA;  60063 MW;  C3DE5E900CB338C4 CRC64;

  Query Match             33.0%;  Score 936.5;  DB 4;  Length 534;
  Best Local Similarity   37.8%;  Pred. No. 7.2e-65;
  Matches  217;  Conservative   78;  Mismatches  168;  Indels  111;  Gaps   12;

Qy     20 PLGPLSPGAL-----------------PRPA------QAQDVVDLDFFTQEPLHLVSPS 55
          |   |:|||                  | |       : : :: ||  |: |:  |: :
Db     18 PPACLAPGALYLALLLHLSLSSQAGDRRPLPVDRAAGLKEKTLILLDVSTKNPVRTVNEN 77

Qy     56 FLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIF----DPKKEST 111
          |||: :| ::  |  :|  | | :| ||||||||||:||||| :|||| |    :| |
Db     78 FLSLQLDPSIIHD-GWLDFLSSKRLVTLARGLSPAFLRFGGKRTDFLQFQNLRNPAKSRG 136

Qy    112 FEERSYWQSQVNQDI------------CKYGSIPPDVEEKLRLEWPYQEQL-LLREHYQK 158
            |:      ||            ||     |||  :|: |   |  | ||:| :
Db    137 GPGPDYYLKNYEDDIVRSDVALDKQKGCKIAQ-HPDVMLELQREKAAQMHLVLLKEQF-- 193

Qy    159 KFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNI 218
          ::|||                  | :|
```

```
Db    194 ---SNTYS----------------------NLIL---------------------------- 202

Qy    219 SWELGNEPNSFLKKADIFINGSQLGEDYIQLHKLLRK-STFKNAKLYGPDVGQPRRKTAK 277
Db    203 -----TEPNNYRTMHGRAVNGSQLGKDYIQLKSLLQPIRIYSRASLYGPNIGRPRKNVIA 257

Qy    278 MLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGK 337
Db    258 LLDGFMKVAGSTVDAVTWQHCYIDGRVVKVMDFLKTRLLDTLSDQIRKIQKVVNTYTPGK 317

Qy    338 KVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDEN 397
Db    318 KIWLEGVVTTSAGGTNNLSDSYAAGFLWLNTLGMLANQGIDVVIRHSFFDHGYNHLVDQN 377

Qy    398 FDPLPDYWLSLLFKKLVGTKVLMASVQGSKRR---------KLRVYLHCTNTDNPRYKEG 448
Db    378 FNPLPDYWLSLLYKRLIGPKVLAVHVAGLQRKPRPGRVIRDKLRIYAHCTNHHNHNYVRG 437

Qy    449 DLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTLKMVDDQTL 508
Db    438 SITLFIINLHRSRKKIKLAGTLRDKLVHQYLLQPYGQEGLKSKSVQLNGQPLVMVDDGTL 497

Qy    509 PPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAAC 542
Db    498 PELKPRPLRAGRTLVIPPVTMGFFVVKNVNALAC 531
```

RESULT 10
Q9HB39
ID Q9HB39 PRELIMINARY; PRT; 480 AA.
AC Q9HB39;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Heparanase-like protein HPA2a.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20483645; PubMed=11027606;
RA McKenzie E., Tyson K., Stamps A., Smith P., Turner P., Barry R.,
RA Hircock M., Patel S., Barry E., Stubberfield C., Terrett J., Page M.;
RT "Cloning and Expression Profiling of Hpa2, a Novel Mammalian
RT Heparanase Family Member.";
RL Biochem. Biophys. Res. Commun. 276:1170-1177(2000).
DR EMBL; AF282885; AAG23421.1; -.
DR InterPro; IPR005199; Glyco_hydro_79n.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
SQ SEQUENCE 480 AA; 53900 MW; F75F89F67AC1FF83 CRC64;

Query Match 31.6%; Score 897.5; DB 4; Length 480;
Best Local Similarity 36.0%; Pred. No. 7e-62;
Matches 202; Conservative 74; Mismatches 146; Indels 139; Gaps 9;

```
Qy     20 PLGPLSPGAL----------------PRPA------QAQDVVDLDFFTQEPLHLVSPS 55
Db     18 PPACLAPGALYLALLLHLSLSSQAGDRRPLPVDRAAGLKEKTLILLDVSTKNPVRTVNEN 77

Qy     56 FLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIF----DPKKEST 111
Db     78 FLSLQLDPSIIHD-GWLDFLSSKRLVTLARGLSPAFLRFGGKRTDFLQFQNLRNPAKSR- 135

Qy    112 FEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVD 171
Db    136 ----------------GGPGPD----------YYLKNYE---------- 148

Qy    172 VLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLK 231
Db    149 ------------------------------------------------DEPNNYRT 156

Qy    232 KADIFINGSQLGEDYIQLHKLLRK-STFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVI 290
```

```
Db    157 MHGRAVNGSQLGKDYIQLKSLLQPIRIYSRASLYGPNIGRPRKNVIALLDGFMKVAGSTV 216

Qy    291 DSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGG 350
Db    217 DAVTWQHCYIDGRVVKVMDFLKTRLLDTLSDQIRKIQKVVNTYTPGKKIWLEGVVTTSAG 276

Qy    351 GAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLF 410
Db    277 GTNNLSDSYAAGFLWLNTLGMLANQGIDVVIRHSFFDHGYNHLVDQNFNPLPDYWLSLLY 336

Qy    411 KKLVGTKVLMASVQGSKRR---------KLRVYLHCTNTDNPRYKEGDLTLYAINLHNVT 461
Db    337 KRLIGPKVLAVHVAGLQRKPRPGRVIRDKLRIYAHCTNHHNHNYVRGSITLFIINLHRSR 396

Qy    462 KYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSS 521
Db    397 KKIKLAGTLRDKLVHQYLLQPYGQEGLKSKSVQLNGQPLVMVDDGTLPELKPRPLRAGRT 456

Qy    522 LGLPAFSYSFFVIRNAKVAAC 542
Db    457 LVIPPVTMGFFVVKNVNALAC 477
```

RESULT 11
Q8T108
ID Q8T108 PRELIMINARY; PRT; 515 AA.
AC Q8T108;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Heparanase-like protein.
GN BMHEPA.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P50; TISSUE=POSTERIOR SILK GLAND;
RA Koike Y., Simada T., Suzuki M.G., Mita K., Abe H., Maeda S.,
RA Osoegawa K., DeJong P.J.;
RT "Genomic sequence of 320kb containing a kettin orthologue on the Z
RT chromosome in Bombyx mori.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB079860; BAB85191.1; -.
SQ SEQUENCE 515 AA; 59769 MW; FB8100ABE6EDDADB CRC64;

Query Match 24.5%; Score 696; DB 5; Length 515;
Best Local Similarity 35.1%; Pred. No. 4.8e-46;
Matches 183; Conservative 83; Mismatches 182; Indels 74; Gaps 18;

```
Qy     46 QEPLHLVSPSFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFD 105
Db     42 QEDIKLISEDFLSFGID-TIEIENYNRINYSDTRLRELAAALSPARLRLGGTMSERLIF- 99

Qy    106 PKKESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTY 165
Db    100 -SKENI---------PISCHNCSYKSYPKSLCQ--LIEKPC-------KHKHKFLPFFIM 140

Qy    166 SRSSVDVLYTFANCSGLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNE 225
Db    141 TGNEWNQINDFCRKTNLKLLFSLNAMLRD-NHGWNEKNARELIEFSKHKQYAIDWQLGNE 199

Qy    226 PNSFLKKADIFINGSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQP---RRKTAKMLKSF 282
Db    200 PNSFQHVFNESVTPQILAKDFEKLRKLLNHNGYRHSLIVGPDTTRPQPHRPECLKYMIEF 259

Qy    283 LKAGGEVIDSVTWHHYYLNGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKV--W 340
Db    260 LGNGSHYINVRSWHQYYLNSKTAKLEDFWNPETFDLL---RQQIETMQNQTKKYKNIPMW 316
```

```
Qy   341 LGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDP 400
         | ||||:|||||| ||:|:|   :|:|||||||:  |  |:|| | | | | |||||  |
Db   317 LSETSSSYGGGAPGLSNTYAGSPLWIDKLGLSAKYNISTVIRQSFIG-GYYSLVDENLKP 375

Qy   401 LPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKE--GDLTLYAIN-- 456
         |||:|:|:|:||||| |||   || :  |  |:|:||||    :|      :||| :|
Db   376 LPDWWISVLYKKLVGNKVL--QVQCNCSRFQRLYIHCTNR---KYTNDTSAVTLYGVNLE 430

Qy   457 ------------LHN----VTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSVQLNGLTL 500
                       ||     : :|: :  | :|::               ||:: ||| |
Db   431 MAKARFFLNGTALHGDDLIIHEYI-ISAPSNNRK--------------SKTILLNGWPL 474

Qy   501 KMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAAC 542
              :  |  |      | |  : || :|  |:||:   :  |
Db   475 YY--ESNLHNLRPNIHRYGRYVSLPPYSIGFWVIKKTSITVC 514
```

RESULT 12
Q9SDA1
ID Q9SDA1 PRELIMINARY; PRT; 521 AA.
AC Q9SDA1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 57.8 kDa protein.
GN F13G24.30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,
RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL133421; CAB62595.1; -.
DR InterPro; IPR005199; Glyco_hydro_79n.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 521 AA; 57831 MW; 07D8664A4B305CC2 CRC64;

Query Match 14.6%; Score 416; DB 10; Length 521;
Best Local Similarity 29.2%; Pred. No. 4.2e-24;
Matches 154; Conservative 68; Mismatches 184; Indels 122; Gaps 24;

```
Qy    75 LGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIPP 134
         |  | |    :   |  :| ||:  | :|:|     |               |:
Db    55 LTRPLLTKAIKAFKPLRIRIGGSLQDQVIYDVGNLKT--------------PCR------ 94

Qy   135 DVEEKLRLEWPYQEQLLLREHYQKKFKNS---TYSRSSV-----DVLYTFANCSGLDLIF 186
                |:|:             ||    :|:  :     | | :|   :|  : |
Db    95 ----------PFQKM-----------NSGLFGFSKGCLHMKRWDELNSFLTATGAVVTF 132

Qy   187 GLNALLRTADLQ-------WNSSNAQLLLDYCSSKGYNI-SWELGNEPNSFLKKADIFIN 238
         |||||     |:       |:  | |  |:|  |||| | ||| ||| :        ::
Db   133 GLNALRGRHKLRGKAWGGAWDHINTQDFLNYTVSKGYVIDSWEFGNELSG--SGVGASVS 190

Qy   239 GSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQP-----RRKTAKMLKSFLKAGGEVIDSV 293
              |:| | |  :: |  :||: |: | :  |      ::   |:|:    :|  |:| |
Db   191 AELYGKDLIVLKDVINK-VYKNSWLHKPILVAPGGFYEQQWYTKLLEI---SGPSVVDVV 246

Qy   294 TWHHYYLNGRT--ATREDFLNPDVLDIFISSVQKVF----QVVESTRPGKKVWLGETSSA 347
         | | | |      |  :  ::|     ::| | | |    | ::   |    |:||:  |
Db   247 THHIYNLGSGNDPALVKKIMDPS----YLSQVSKTFKDVNQTIQEHGPWASPWVGESGGA 302

Qy   348 YGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDE-NFDPLPDYWL 406
         | |   :|||| | :||:||:|||   :|  ||   | | | |:::  | | |||:
Db   303 YNSGGRHVSDTFIDSFWYLDQLGMSARHNTKVYCRQTLVG-GFYGLLEKGTFVPNPDYYS 361

Qy   407 SLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYL-- 464
         :||: :|:|  ||      |   :|||| ||:           :||  ||| | : :
Db   362 ALLWHRLMGKGVLAVQTDGPP--QLRVYAHCSK------GRAGVTLLLINLSNQSDFTVS 413

Qy   465 ---------------------RLPYPFS---NKQVDKYLLRP---LGPHG--LLSKSVQL 495
                              |  |||   :|  | || |    | |    | ||:: |
Db   414 VSNGINVVLNAESRKKKSLLDTLKRPFSWIGSKASDGYLNREEYHLTPENGVLRSKTMVL 473

Qy   496 NGLTLKMVDDQTLPPLMEKPLRP-GSSLGLPAFSYSFFVIRNAKVAAC 542
         || :||      :| | |  ||   | | :   | || |: |   :||
Db   474 NGKSLKPTATGDIPSL-EPVLRSVNSPLNVLPLSMSFIVLPNFDASAC 520

RESULT 13
Q9FF10
ID Q9FF10 PRELIMINARY; PRT; 543 AA.
AC Q9FF10;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similarity to heparanase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT P1 clones.";
RL DNA Res. 4:215-230(1997).
DR EMBL; AB005249; BAB09947.1; -.
DR InterPro; IPR005199; Glyco_hydro_79n.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF03662; Glyco_hydro_79n; 1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
SQ SEQUENCE 543 AA; 60250 MW; 0FA2248948282FF6 CRC64;

Query Match 14.6%; Score 416; DB 10; Length 543;
Best Local Similarity 29.2%; Pred. No. 4.5e-24;
Matches 154; Conservative 68; Mismatches 184; Indels 122; Gaps 24;

```
Qy   75 LGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIPP 134
         | | | : | :| ||: | :|:| | |:
Db   77 LTRPLLTKAIKAFKPLRIRIGGSLQDQVIYDVGNLKT--------------PCR------ 116

Qy   135 DVEEKLRLEWPYQEQLLLREHYQKKFKNS---TYSRSSV-----DVLYTFANCSGLDLIF 186
                |:|: || :|: : | | :| :| : |
Db   117 ----------PFQKM-----------NSGLFGFSKGCLHMKRWDELNSFLTATGAVVTF 154

Qy   187 GLNALLRTADLQ-------WNSSNAQLLLDYCSSKGYNI-SWELGNEPNSFLKKADIFIN 238
         ||||| |: |: | | |:| |||| | ||| ||| :   ::
Db   155 GLNALRGRHKLRGKAWGGAWDHINTQDFLNYTVSKGYVIDSWEFGNELSG--SGVGASVS 212

Qy   239 GSQLGEDYIQLHKLLRKSTFKNAKLYGPDVGQP-----RRKTAKMLKSFLKAGGEVIDSV 293
              |:| | | :: | :||: |: | : | :: |:|:   :| |:| |
Db   213 AELYGKDLIVLKDVINK-VYKNSWLHKPILVAPGGFYEQQWYTKLLEI---SGPSVVDVV 268

Qy   294 TWHHYYLNGRT--ATREDFLNPDVLDIFISSVQKVF----QVVESTRPGKKVWLGETSSA 347
         | | | | | : ::| ::| | | | | :: | |:||: |
Db   269 THHIYNLGSGNDPALVKKIMDPS----YLSQVSKTFKDVNQTIQEHGPWASPWVGESGGA 324

Qy   348 YGGGAPLLSDTFAAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDE-NFDPLPDYWL 406
         | | :|||| | :||:||:||| :| || | | | |::: | | |||:
Db   325 YNSGGRHVSDTFIDSFWYLDQLGMSARHNTKVYCRQTLVG-GFYGLLEKGTFVPNPDYYS 383
```
```

```
Qy    407 SLLFKKLVGTKVLMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYL-- 464
          :||: :|:|  ||     |    :|||| ||:          :||  ||| | : :
Db    384 ALLWHRLMGKGVLAVQTDGPP--QLRVYAHCSK------GRAGVTLLLINLSNQSDFTVS 435

Qy    465 --------------------RLPYPFS---NKQVDKYLLRP---LGPHG--LLSKSVQL 495
                               |  |||   :|  | || |     | |    | ||:: |
Db    436 VSNGINVVLNAESRKKKSLLDTLKRPFSWIGSKASDGYLNREEYHLTPENGVLRSKTMVL 495

Qy    496 NGLTLKMVDDQTLPPLMEKPLRP-GSSLGLPAFSYSFFVIRNAKVAAC 542
          || :||      :| | |  ||   | | :   | || |: |   :||
Db    496 NGKSLKPTATGDIPSL-EPVLRSVNSPLNVLPLSMSFIVLPNFDASAC 542


RESULT 14
Q9LRC8
ID   Q9LRC8      PRELIMINARY;      PRT;   527 AA.
AC   Q9LRC8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta-glucuronidase.
GN   SGUS.
OS   Scutellaria baicalensis.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Lamiales; Lamiaceae; Scutellaria.
OX   NCBI_TaxID=65409;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20418130; PubMed=10858442;
RA   Sasaki K., Taura F., Shoyama Y., Morimoto S.;
RT   "Molecular Characterization of a Novel beta-Glucuronidase from
RT   Scutellaria baicalensis Georgi.";
RL   J. Biol. Chem. 275:27466-27472(2000).
DR   EMBL; AB040072; BAA97804.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ   SEQUENCE   527 AA;  58772 MW;  A5DE7C423F2A1E2B CRC64;

  Query Match             13.4%;  Score 381;  DB 10;  Length 527;
  Best Local Similarity   25.2%;  Pred. No. 2.4e-21;
  Matches  133;  Conservative   83;  Mismatches  178;  Indels  134;  Gaps   22;

Qy     55 SFLSVTIDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEE 114
          |||:: :: |:              :|   :  :|  ||||||  | |::   ::   :
Db     69 SFLNLDLNNNI--------------IRNAVKEFAPLKLRFGGTLQDRLVYQTSRDEPCDS 114

Qy    115 RSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLY 174
            |  :  |                                | ::  |    | :
Db    115 TFYNNTNLILD---------------------------------FSHACLSLDRWDEIN 140

Qy    175 TFANCSGLDLIFGLNAL------------------LRTADLQWNSSNAQLLLDYCSSKG 215
          |   :| : :||||||                  ||  :|: ||:: |::|   ||
Db    141 QFILETGSEAVFGLNALRGKTVEIKGIIKDGQYLGETTTAVGEWDYSNSKFLIEYSLKKG 200

Qy    216 YN--ISWELGNEPNSFLKKADIFINGSQLGEDYI----QLHKLLRKSTFKNAKLYGPDVG 269
          |     | ||||    |    :||  |   |||     :||:|::       ::|
Db    201 YKHIRGWTLGNE----LGGHTLFIGVSP--EDYANDAKKLHELVK-------EIYQDQGT 247

Qy    270 QPRRKTAKMLKSFLKAGGEVID-------------SVTWHHYY---LNGRTATREDFLN 312
          |           : | | : |               |  || |     |  | ::  |
Db    248 MP----------LIIAPGAIFDLEWYTEFIDRTPELHVATHHMYNLGSGGDDALKDVLLT 297

Qy    313 PDVLDIFISSVQKVFQVVESTRPGKK--VWLGETSSAYGGGAPLLSDTFAAGFMWLDKLG 370
          |    |: :  | : : ||| |   |:||   |:  |   :|:||  || :|: ||
Db    298 ASFFDEATKSMYEGLQKIVN-RPGTKAVAWIGEAGGAFNSGQDGISNTFINGFWYLNMLG 356

Qy    371 LSARMGIEVVMRQVFFGAGNYHLVDE-NFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRR 429
          || :  :   ||   | ||| |:    : | |||: :||: :|:|:|:|||   : |:|
Db    357 YSALLDTKTFCRQTLTG-GNYGLLQTGTYIPNPDYYSALLWHRLMGSKVLKTEIVGTK-- 413

Qy    430 KLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRL---PYPFSNKQVDKYLLRPLGPH 486
          : :| ||        |   :|:  :| |:    :::   |  : :|: ::| | |:  :
Db    414 NVYIYAHCAK------KSNGITMLVLN-HDGESSVKISLDPSKYGSKR-EEYHLTPVN-N 464

Qy    487 GLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVI 534
          | |: |:|||  | :     :| |   |:   :|  |    |||  :
Db    465 NLQSRLVKLNGELLHLDPSGVIPAL--NPVEKDNSKQLEVAPYSFMFV 510


RESULT 15
Q9FZP1
ID   Q9FZP1      PRELIMINARY;      PRT;   536 AA.
AC   Q9FZP1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similarity to heparanase.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RA   Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB028613; BAB10787.1; -.
DR   InterPro; IPR005199; Glyco_hydro_79n.
DR   Pfam; PF03662; Glyco_hydro_79n; 1.
DR   PRINTS; PR01656; VACCYTOTOXIN.
SQ   SEQUENCE   536 AA;  59654 MW;  24202B8E82F3DB0E CRC64;

  Query Match             12.8%;  Score 363;  DB 10;  Length 536;
  Best Local Similarity   24.6%;  Pred. No. 6.2e-20;
  Matches  127;  Conservative   81;  Mismatches  176;  Indels  132;  Gaps   20;

Qy     88 SPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIPPDVEEKLRLEWPYQ 147
          :|  :| ||| | :|::                                || ::      |:
Db     91 APLKIRIGGTLQDIVIYE--------------------------TPDSKQPC---LPFT 120

Qy    148 EQLLLREHYQKKFKNST----YSRSSV-----DVLYTFANCSGLDLIFGLNALL------ 192
                 |||:   .|:: :     | |  |   :|  :|||||||
Db    121 -------------KNSSILFGYTQGCLPMRRWDELNAFFRKTGTKVIFGLNALSGRSIKS 167

Qy    193 -RTADLQWNSSNAQLLLDYCSSKGYNI-SWELGNEPNSFLKKADIFINGSQLGE------ 244
           |   || :||:  : : :   | |  ||||||           : || :|
Db    168 NGEAIGAWNYTNAESFIRFTAENNYTIDGWELGNE--------LCGSGVGARVGANQ 216

Qy    245 ---DYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLN 301
             | | |  :: :  :||       :|            :|      ::: | | | |
Db    217 YAIDTINLRNIVNR-VYKNVSPMPLVIGPGGFFEVDWFTEYLNKAENSLNATTRHIYDLG 275

Qy    302 GRTATR--EDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTF 359
             |  ||| ||    | : :  :::::      |:||:  ||  |  |:|: |
Db    276 PGVDEHLIEKILNPSYLDQEAKSFRSLKNIIKNSSTKAVAWVGESGGAYNSGRNLVSNAF 335

Qy    360 AAGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVD-ENFDPLPDYWLSLLFKKLVGTKV 418
          | :||:||:::    :   ||   | ||| |::  || | ||| |::|:::::|:| |
Db    336 VYSFWYLDQLGMASLYDTKTYCRQSLIG-GNYGLLNTTNFTPNPDYYSALIWRQLMGRKA 394

Qy    419 LMASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNV----------------T 461
          |  :  |:|  |:| | ||      |  :| :|:  :|| |                  |
Db    395 LFTTFSGTK--KIRSYTHCA-----RQSKG-ITVLLMNLDNTTTVVAKVELNNSFSLRHT 446

Qy    462 KYLRLPYPFSNKQVDKYLLRPLGPHGLL----------SKSVQLNGLTLKMVDDQ 506
          |:::    | |:    |   ||:|::            |::: |||  |::
Db    447 KHMK-----SYKRASSQLFG--GPNGVIQREEYHLTAKDGNLHSQTMLLNGNALQVNSMG 499

Qy    507 TLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAAC 542
```

```
          |||:    :     : :  :|  |  :||  | ||
Db    500 DLPPIEPIHINSTEPITIAPYSIVFVHMRNVVVPAC 535
```

Search completed: November 20, 2002, 11:37:17
Job time : 40 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: November 20, 2002, 11:31:59 ; Search time 39 Seconds
(without alignments)
1855.259 Million cell updates/sec

Title: US-09-759-207-2
Perfect score: 2842
Sequence: 1 MLLRSKPALPPPLMLLLLGP.....LPAFSYSFFVIRNAKVAACI 543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2842	100.0	543	20	AAY02345	A human heparanase
2	2842	100.0	543	21	AAB08849	Amino acid sequenc
3	2842	100.0	543	21	AAY57590	Human heparanase.
4	2842	100.0	543	21	AAY52990	Human heparanase p
5	2842	100.0	543	22	AAY97635	Human heparanase p
6	2842	100.0	543	23	ABB07813	Human heparanase s
7	2842	100.0	592	20	AAY02346	A human heparanase
8	2842	100.0	592	21	AAB08850	Amino acid sequenc
9	2838	99.9	543	20	AAY17082	Human heparanase e
10	2838	99.9	543	22	AAB86206	Human heparanase i
11	2838	99.9	588	20	AAY30124	A human protein wi
12	2826	99.4	543	22	AAB88361	Human membrane or
13	2764	97.3	530	20	AAY34173	Human pre-prohepar
14	2737	96.3	532	20	AAY17083	Seq ID No: 15 of W
15	2673.5	94.1	527	23	ABB07815	Chicken signal pep
16	2146	75.5	535	21	AAB08851	A murine heparanas
17	2146	75.5	535	23	ABB07811	Mouse heparanase s
18	2123	74.7	536	23	ABB07812	Rat heparanase seq
19	1645.5	57.9	523	23	ABB07814	Chicken heparanase
20	1614	56.8	380	20	AAY17085	Rat heparanase enz
21	1602	56.4	380	20	AAY17084	Mouse heparanase e
22	1154.5	40.6	592	22	AAU07424	Human heparanase-1
23	1154.5	40.6	592	22	AAY97632	Human heparanase,
24	1148.5	40.4	592	22	AAB81062	Human Heparanase-2
25	1147.5	40.4	592	22	AAB85215	Heparanase-like pr
26	1142.5	40.2	582	23	AAE18326	Human heparanase-2
27	1112.5	39.1	538	22	AAY97633	Human heparanase,
28	1106.5	38.9	528	23	AAE18327	Human heparanase-2
29	936.5	33.0	534	22	AAB85216	Heparanase-like pr
30	936.5	33.0	534	23	AAM50337	Human prepro-hepar
31	927.5	32.6	492	22	AAB84664	Amino acid sequenc
32	897.5	31.6	480	22	AAU07418	Novel human extrac
33	897.5	31.6	480	22	AAB85217	Heparanase-like pr
34	897.5	31.6	480	22	AAY97634	Human heparanase,
35	892.5	31.4	470	23	AAE18328	Human heparanase-2
36	891.5	31.4	439	22	AAU07423	Human heparanase-1
37	788	27.7	331	23	AAM50383	Human heparanase I
38	663	23.3	488	22	AAB31469	Amino acid sequenc
39	645	22.7	488	22	AAB31470	Amino acid sequenc
40	642	22.6	488	22	AAB31472	Amino acid sequenc
41	622	21.9	488	22	AAB31471	Amino acid sequenc
42	528.5	18.6	214	22	AAM99905	Human excretory re
43	528.5	18.6	214	22	AAM43704	Human bladder anti
44	338.5	11.9	156	22	AAG65963	Human heparanase-1
45	277.5	9.8	256	21	AAG13479	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAY02345
ID AAY02345 standard; Protein; 543 AA.
XX
AC AAY02345;
XX
DT 09-JUL-1999 (first entry)
XX
DE A human heparanase protein.
XX
KW Heparanase; hp; modulator; heparin-binding growth factor;
KW cellular response; cytokine; cell interaction; plasma lipoprotein;
KW cellular susceptibility; infection; disintegration;
KW neurodegenerative plaque; wound healing; angiogenesis; restenosis;
KW athersclerosis; inflammation; neurodegenerative disease; neutralise;
KW plasma heparin; micrometastasis; autoimmune lesion; renal failure.
XX
OS Homo sapiens.
XX
PN WO9911798-A1.
XX
PD 11-MAR-1999.
XX
PF 31-AUG-1998; 98WO-US17954.
XX
PR 02-JUL-1998; 98US-0109386.
PR 02-SEP-1997; 97US-0922170.
XX
PA (FRIE/) FRIEDMAN M M.
PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
PA (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX
PI Feinstein E, Pecker I, Vlodavsky I;

```
XX
DR   WPI; 1999-302255/25.
DR   N-PSDB; AAX35648.
XX
PT   New human polynucleotide useful for treating angiogenesis,
PT   restenosis, and inflammation
XX
PS   Claim 6; Fig 1; 63pp; English.
XX
CC   The specification describes a polypeptide having heparanase (hp)
CC   activity. The recombinant protein is used as a modulator of
CC   heparin-binding growth factors, cellular responses to heparin-binding
CC   growth factors and cytokines, cell interaction with plasma lipoproteins,
CC   cellular susceptibility to viral, protozoal and bacterial infections
CC   or disintegration of neurodegenerative plaques. Heparanase may be
CC   useful for conditions such as wound healing, angiogenesis, restenosis,
CC   atherosclerosis, inflammation, neurodegenerative diseases, and viral
CC   infections. Mammalian heparanase can be used to neutralize plasma
CC   heparin, and anti-heparanase antibodies may be applied for
CC   immunodetection and diagnosis of micrometastases, autoimmune lesions,
CC   and renal failure in biopsy specimens, plasma samples, and body fluids.
CC   The present sequence represents human heparanase.
XX
SQ   Sequence   543 AA;

  Query Match             100.0%;  Score 2842;  DB 20;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy     541 ACI 543
           |||
Db     541 ACI 543


RESULT 2
AAB08849
ID   AAB08849 standard; Protein; 543 AA.
XX
AC   AAB08849;
XX
DT   15-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a human heparanase polypeptide.
XX
KW   Human; heparanase; gene therapy; tumour; inflammation; autoimmunity;
KW   heparin-binding growth factor; cytokine; neurodegenerative plaque;
KW   wound healing; infection; burn; angiogenesis; restenosis;
KW   atherosclerosis; inflammation; neurodegenerative disease;
KW   Gerstmann-Straussler Syndrome; Creutzfeldt-Jakob disease.
XX
OS   Homo sapiens.
XX
PN   WO200052178-A1.
XX
PD   08-SEP-2000.
XX
PF   14-FEB-2000; 2000WO-US03542.
XX
PR   01-MAR-1999;   99US-0258892.
XX
PA   (INSI-) INSIGHT STRATEGY & MARKETING LTD.
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
PA   (FRIE/) FRIEDMAN M M.
XX
PI   Pecker I,  Vlodavsky I,  Feinstein E;
XX
DR   WPI; 2000-579289/54.
DR   N-PSDB; AAA75051.
XX
PT   New polynucleotides encoding a polypeptide having heparanase activity,
PT   useful in wound healing and in gene therapy, particularly in treating
PT   tumour, inflammation, autoimmunity, neurodegenerative diseases
XX
PS   Claim 22; Fig 1; 152pp; English.
XX
CC   The present sequence represents a human protein with heparanase catalytic
CC   activity. The heparanase (hpa) polynucleotide is useful in gene therapy,
CC   particularly in treating tumour, inflammation or autoimmunity.
CC   Particularly, the polynucleotide is useful in modulating the
CC   bioavailability of heparin-binding growth factors, cellular responses
CC   to heparin-binding growth factors (e.g. bFGF) and cytokines
CC   (e.g. interleukin (IL)-8), cell interaction with plasma lipoproteins,
CC   cellular susceptibility to certain viral and some bacterial and protozoa
CC   infections, or disintegration of neurodegenerative plaques. The
CC   polynucleotide is also useful in wound healing (e.g. thermal, chemical
CC   or radiation burns), and in the treatment of angiogenesis, restenosis,
CC   atherosclerosis, inflammation, neurodegenerative diseases (Gerstmann-
CC   Straussler Syndrome or Creutzfeldt-Jakob disease), and some viral,
CC   bacterial or protozoa infections.
XX
SQ   Sequence   543 AA;

  Query Match             100.0%;  Score 2842;  DB 21;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy    541 ACI 543
          |||
Db    541 ACI 543
```

RESULT 3
AAY57590
ID AAY57590 standard; Protein; 543 AA.
XX
AC AAY57590;
XX
DT 02-MAR-2000 (first entry)
XX
DE Human heparanase.
XX
KW Human; heparanase; hpa; genetic modification; expression; anticancer;
KW angiogenesis; anti-angiogenic; antiproliferative; antiviral; antitumour;
KW anti-atherosclerotic; anti-inflammatory; antineurodegeneration;
KW heparan sulphate; heparin-binding growth factor; tumour angiogenesis;
KW metastasis; wound healing; restenosis; atherosclerosis; inflammation;
KW neurodegeneration; viral infection; cystic fibrosis; cancer; diagnosis;
KW micrometastasis; autoimmune lesion; kidney failure.
XX
OS Homo sapiens.
XX
PN WO9957244-A1.
XX
PD 11-NOV-1999.
XX
PF 29-APR-1999; 99WO-US09256.
XX
PR 01-MAY-1998; 98US-0071618.
PR 02-MAR-1999; 99US-0260038.
XX
PA (INSI-) INSIGHT STRATEGY & MARKETING LTD.
PA (FRIE/) FRIEDMAN M M.
XX
PI Ben-Artzi H, Ayal-Hershkovitz M, Yacoby-Zeevi O, Pecker I, Peleg Y;
PI Shlomi Y;
XX
DR WPI; 2000-062144/05.
DR N-PSDB; AAZ39195.
XX
PT Engineered cells that express recombinant heparanase, useful
PT therapeutically, e.g. for treating angiogenesis and to screen for
PT specific inhibitors, potential anticancer agents -
XX
PS Claim 3; Page 107-109; 118pp; English.
XX
CC The present invention describes genetically modified cells (A) containing
CC a polynucleotide (I) that encodes a polypeptide with heparanase activity,
CC and express recombinant heparanase (II). Heparanase cleaves heparan
CC sulphate (HS) at specific intrachain sites, resulting in release of
CC heparin-binding growth factors, enzymes and proteins that are sequestered
CC by HS in basement membranes, extracellular matrix or cell surfaces. It
CC may also be implicated in tumour angiogenesis and metastases. (II) is
CC potentially useful in wound healing and for treating angiogenesis,
CC restenosis, atherosclerosis, inflammation, neurodegeneration, viral
CC infection and cystic fibrosis. It can also be used to neutralise heparin
CC (an alternative to protamine) and to screen for specific inhibitors
CC (potentially useful for treating cancer and metastases). Antibodies
CC raised against (II) are used for immunodetection and diagnosis of
CC micrometastases, autoimmune lesions and kidney failure. (A) provide (II)
CC in large quantities, in a form that is homogeneously processed and
CC activated/neutralised by a dedicated protease. The present sequence
CC represents human heparanase.
XX
SQ Sequence 543 AA;

```
  Query Match             100.0%;  Score 2842;  DB 21;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy     61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy    121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy    181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy    541 ACI 543
          |||
Db    541 ACI 543
```

RESULT 4
AAY52990
ID AAY52990 standard; Protein; 543 AA.
XX
AC AAY52990;
XX
DT 21-FEB-2000 (first entry)
XX
DE Human heparanase protein sequence.
XX

```
KW   Human; heparanase; hpa; diagnosis; therapy; tumour; cytostatic;
KW   antidiabetic; immunomodulatory; anti-inflammatory; nephrotropic;
KW   metastasis; adenocarcinoma; squamous cell carcinoma; teratocarcinoma;
KW   mesothelioma; melanoma; lymphoma; leukemia; cancer; sepsis; diabetes;
KW   inflammation; haemorrhagic nephritis; nephrotic syndrome;
KW   autoimmune disease; anticancer; kidney disease.
XX
OS   Homo sapiens.
XX
PN   WO9957153-A1.
XX
PD   11-NOV-1999.
XX
PF   29-APR-1999;   99WO-US09255.
XX
PR   01-MAY-1998;   98US-0071739.
XX
PA   (INSI-) INSIGHT STRATEGY & MARKETING LTD.
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
PA   (FRIE/) FRIEDMAN M M.
XX
PI   Pecker I, Vlodavsky I, Friedman Y, Perets T;
XX
DR   WPI; 2000-052944/04.
DR   N-PSDB; AAZ33290.
XX
PT   Heparanase-specific molecular probes useful for diagnosis and
PT   treatment, e.g. of tumors, and for targeted drug delivery -
XX
PS   Example; Page 81-82; 90pp; English.
XX
CC   The present invention describes heparanase-specific molecular probes,
CC   useful for methods of detecting heparanase in situ. The probes and
CC   anti-heparanase antibodies are used to detect or quantify the expression
CC   of heparanase, for diagnosis and monitoring of diseases (especially
CC   metastasis), for treatment of heparanase-associated diseases (e.g.
CC   tumours, (adeno)carcinoma, squamous cell carcinoma, teratocarcinoma,
CC   mesothelioma, melanoma, lymphoma or leukemia, a solid cancer (or its
CC   metastases) derived from liver, prostate, bladder, breast, ovary,
CC   cervix, colon, skin, intestine, stomach, uterus and pancreas, kidney
CC   disease, diabetes and inflammation, haemorrhagic nephritis, nephrotic
CC   syndrome, sepsis and inflammatory or autoimmune disease), for targeted
CC   drug delivery (e.g. of anticancer agents) and as research reagents.
CC   The present sequence represents human heparanase, which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   543 AA;

  Query Match            100.0%;  Score 2842;  DB 21;  Length 543;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy     541 ACI 543
           |||
Db     541 ACI 543


RESULT 5
AAY97635
ID   AAY97635 standard; Protein; 543 AA.
XX
AC   AAY97635;
XX
DT   20-APR-2001  (first entry)
XX
DE   Human heparanase protein sequence.
XX
KW   Heparanase; hnhp1; wound healing; angiogenesis; restenosis; Scrape;
KW   atherosclerosis; inflammation; pulmonary disease; Alzheimer's disease;
KW   neurodegenerative disease; Creutzfeldt-Jakob disease; viral infection;
KW   gene therapy; human.
XX
OS   Homo sapiens.
XX
PN   WO200100643-A2.
XX
PD   04-JAN-2001.
XX
PF   19-JUN-2000; 2000WO-IL00358.
XX
PR   25-JUN-1999;   99US-0140801.
XX
PA   (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX
PI   Pecker I, Michal I, Itzhaki H;
XX
DR   WPI; 2001-137930/14.
XX
PT   New polynucleotides and polypeptides that are distantly homologous to
PT   heparanase, useful in wound healing, as well as in gene therapy
PT   protocols for angiogenesis, restenosis, atherosclerosis, or
PT   inflammation -
XX
PS   Disclosure; Page 64-65; 67pp; English.
XX
CC   This sequence represents a heparanase of the invention.
CC   The heparanase DNA and protein sequences are useful in wound healing,
CC   angiogenesis, restenosis, atherosclerosis, inflammation, pulmonary
CC   diseases, neurodegenerative diseases (such as Scrape, Alzheimer's
CC   disease, and Creutzfeldt-Jakob disease) or viral infections. The
CC   heparanase coding sequence is particularly useful in gene therapy.
XX
SQ   Sequence   543 AA;

  Query Match            100.0%;  Score 2842;  DB 22;  Length 543;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy     61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy    121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy    181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy    541 ACI 543
          |||
Db    541 ACI 543


RESULT 6
ABB07813
ID   ABB07813 standard; protein; 543 AA.
XX
AC   ABB07813;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Human heparanase sequence.
XX
KW   Heparanase; catalytic; cytostatic; antiviral; antibacterial; enzyme;
KW   anti-protozoan; neuroprotective; heparin; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..35
FT                   /note= "signal peptide"
FT   Protein         36..543
FT                   /note= "mature protein"
XX
PN   US2002034810-A1.
XX
PD   21-MAR-2002.
XX
PF   16-AUG-2001; 2001US-0930218.
XX
PR   20-SEP-2000; 2000US-0666390.
XX
PA   (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX
PI   Goldshmidt O,  Pecker I,  Vlodavsky I,  Michal I,  Zcharia E;
XX
DR   WPI; 2002-338926/37.
XX
PT   Nucleic acid encoding avian and reptile heparanase polypeptide is
PT   useful to treat various heparin-related disorders and the signal
PT   peptide is useful in production of membrane-targeted or secreted
PT   recombinant proteins
XX
PS   Disclosure; Fig 1a; 39pp; English.
XX
CC   The invention relates to an isolated avian and reptile nucleic acid,
CC   encoding a polypeptide with heparanase catalytic activity. The signal
CC   peptide of the nucleic acid can be used to express membrane-associated or
CC   secreted proteins in heterologous expression systems. The encoded
CC   polypeptides can be used to prevent tumour angiogenesis, metastasis and
CC   invasion, and to intervene with pathologies associated with impaired
CC   heparin-binding growth factors, cellular responses to heparin-binding
CC   growth factors and cytokines, cell interaction with plasma lipoproteins,
CC   cellular susceptibility to viral, protozoa and bacterial infections or
CC   disintegration of neurodegenerative plaques. The present sequence
CC   represents a human heparanase protein sequence used in similarity
CC   studies.
XX
SQ   Sequence   543 AA;

  Query Match           100.0%;  Score 2842;  DB 23;  Length 543;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy     61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy    121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy    181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy    541 ACI 543
          |||
Db    541 ACI 543


RESULT 7
AAY02346
ID   AAY02346 standard; Protein; 592 AA.
XX
AC   AAY02346;
```

```
XX
DT   09-JUL-1999  (first entry)
XX
DE   A human heparanase protein.
XX
KW   Heparanase; hpa; modulator; heparin-binding growth factor;
KW   cellular response; cytokine; cell interaction; plasma lipoprotein;
KW   cellular susceptibility; infection; disintegration;
KW   neurodegenerative plaque; wound healing; angiogenesis; restenosis;
KW   atherscIerosis; inflammation; neurodegenerative disease; neutralise;
KW   plasma heparin; micrometastasis; autoimmune lesion; renal failure.
XX
OS   Homo sapiens.
XX
PN   WO9911798-A1.
XX
PD   11-MAR-1999.
XX
PF   31-AUG-1998;   98WO-US17954.
XX
PR   02-JUL-1998;   98US-0109386.
PR   02-SEP-1997;   97US-0922170.
XX
PA   (FRIE/) FRIEDMAN M M.
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
PA   (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX
PI   Feinstein E,  Pecker I,  Vlodavsky I;
XX
DR   WPI; 1999-302255/25.
DR   N-PSDB; AAX35650.
XX
PT   New human polynucleotide useful for treating angiogenesis,
PT   restenosis, and inflammation
XX
PS   Claim 6; Page 65-66; 63pp; English.
XX
CC   The specification describes a polypeptide having heparanase (hpa)
CC   activity. The recombinant protein is used as a modulator of
CC   heparin-binding growth factors, cellular responses to heparin-binding
CC   growth factors and cytokines, cell interaction with plasma lipoproteins,
CC   cellular susceptibility to viral, protozoal and bacterial infections
CC   or disintegration of neurodegenerative plaques. Heparanase may be
CC   useful for conditions such as wound healing, angiogenesis, restenosis,
CC   atherscIerosis, inflammation, neurodegenerative diseases, and viral
CC   infections. Mammalian heparanase can be used to neutralize plasma
CC   heparin, and anti-heparanase antibodies may be applied for
CC   immunodetection and diagnosis of micrometastases, autoimmune lesions,
CC   and renal failure in biopsy specimens, plasma samples, and body fluids.
CC   The present sequence represents human heparanase.
XX
SQ   Sequence   592 AA;

  Query Match             100.0%;  Score 2842;  DB 20;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      50 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLSVT 109

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 169

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 229

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     230 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 289

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     290 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 349

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     350 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 409

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     410 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 469

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     470 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 529

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     530 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 589

Qy     541 ACI 543
           |||
Db     590 ACI 592


RESULT 8
AAB08850
ID   AAB08850 standard; Protein; 592 AA.
XX
AC   AAB08850;
XX
DT   15-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a human heparanase polypeptide.
XX
KW   Human; heparanase; gene therapy; tumour; inflammation; autoimmunity;
KW   heparin-binding growth factor; cytokine; neurodegenerative plaque;
KW   wound healing; infection; burn; angiogenesis; restenosis;
KW   atherosclerosis; inflammation; neurodegenerative disease;
KW   Gerstmann-Straussler Syndrome; Creutzfeldt-Jakob disease.
XX
OS   Homo sapiens.
XX
PN   WO200052178-A1.
XX
PD   08-SEP-2000.
XX
PF   14-FEB-2000; 2000WO-US03542.
XX
PR   01-MAR-1999;   99US-0258892.
XX
PA   (INSI-) INSIGHT STRATEGY & MARKETING LTD.
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
PA   (FRIE/) FRIEDMAN M M.
XX
PI   Pecker I,  Vlodavsky I,  Feinstein E;
XX
DR   WPI; 2000-579289/54.
DR   N-PSDB; AAA75053.
XX
PT   New polynucleotides encoding a polypeptide having heparanase activity,
PT   useful in wound healing and in gene therapy, particularly in treating
PT   tumour, inflammation, autoimmunity, neurodegenerative diseases  -
XX
PS   Claim 22; Page 122-123; 152pp; English.
XX
CC   The present sequence represents a human protein with heparanase catalytic
CC   activity. The heparanase (hpa) polynucleotide is useful in gene therapy,
CC   particularly in treating tumour, inflammation or autoimmunity.
CC   Particularly, the polynucleotide is useful in modulating the
CC   bioavailability of heparin-binding growth factors, cellular responses
CC   to heparin-binding growth factors (e.g. bFGF) and cytokines
CC   (e.g. interleukin (IL)-8), cell interaction with plasma lipoproteins,
```

```
CC   cellular susceptibility to certain viral and some bacterial and protozoa
CC   infections, or disintegration of neurodegenerative plaques. The
CC   polynucleotide is also useful in wound healing (e.g. thermal, chemical
CC   or radiation burns), and in the treatment of angiogenesis, restenosis,
CC   atherosclerosis, inflammation, neurodegenerative diseases (Gerstmann-
CC   Straussler Syndrome or Creutzfeldt-Jakob disease), and some viral,
CC   bacterial or protozoa infections.
XX
SQ   Sequence   592 AA;

  Query Match             100.0%;  Score 2842;  DB 21;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-273;
  Matches  543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      50 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLSVT 109

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 169

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 229

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     230 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 289

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     290 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 349

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     350 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 409

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     410 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 469

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     470 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 529

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     530 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 589

Qy     541 ACI 543
           |||
Db     590 ACI 592


RESULT 9
AAY17082
ID   AAY17082 standard; Protein; 543 AA.
XX
AC   AAY17082;
XX
DT   21-JUL-1999  (first entry)
XX
DE   Human heparanase enzyme.
XX
KW   Heparanase; endoglucuronidase; heparan sulfate proteoglycan; enzyme;
KW   metastasis; angiogenesis; wound healing; angioplasty-induced restenosis;
KW   arteriosclerosis; atherosclerosis; inflammation; tissue development;
KW   human; HSPG.
XX
OS   Homo sapiens.
XX
PN   WO9921975-A1.
XX
PD   06-MAY-1999.
XX
PF   28-OCT-1998;   98WO-AU00898.
XX
PR   09-DEC-1997;   97AU-0000812.
PR   28-OCT-1997;   97AU-0000062.
XX
PA   (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI   Freeman CG,  Hamdorf BJ,  Hulett MD,  Parish CR;
XX
DR   WPI; 1999-312956/26.
DR   N-PSDB; AAX37259.
XX
PT   Polynucleotides encoding mammalian endoglucuronidases, especially
PT   heparanases, useful to promote wound healing
XX
PS   Claim 6; Page 69-73; 112pp; English.
XX
CC   The invention relates to nucleic acid sequences that encode heparanase
CC   enzymes having  endoglucuronidase activity. Recombinant heparanases are
CC   capable of removing the HS side chain from heparan sulfate proteoglycan
CC   (HSPG). Sulfated oligosaccharides, sulphonates or HSPG can be used to
CC   inhibit heparanase, this is useful for treatment of a physiological or
CC   medical condition associated with elevated heparanase activity, such as
CC   metastasis, angiogenesis, wound healing, angioplasty-induced restenosis,
CC   arteriosclerosis, atherosclerosis and inflammation. The human, murine and
CC   rat heparanases can be used to enhance wound healing, especially
CC   associated with tissue development and repair. The conditions mentioned
CC   above can be diagnosed using specific antibodies, and also using primers
CC   and probes specific for the heparanase polynucleotides. Other uses of the
CC   heparanases include sequencing sulfated molecules such as HSPG. The
CC   present sequence represents a human heparanase.
XX
SQ   Sequence   543 AA;

  Query Match             99.9%;  Score 2838;  DB 20;  Length 543;
  Best Local Similarity   99.8%;  Pred. No. 3.3e-273;
  Matches  542;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQADVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
```

```
Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy     541 ACI 543
           |||
Db     541 ACI 543
```

RESULT 10
AAB86206
ID AAB86206 standard; Protein; 543 AA.
XX
AC AAB86206;
XX
DT 24-AUG-2001 (first entry)
XX
DE Human heparanase inhibitor protein.
XX
KW Heparanase; inhibitor; cardiac insufficiency; cardiant; nephrotropic;
KW hepatotropic; veterinary medicine; congestive heart failure; dyspnoea;
KW primary cardiomyopathy; peripheral odema; pulmonary congestion;
KW hepatic congestion; hydrothorax; ascite; nocturia; human.
XX
OS Homo sapiens.
XX
PN DE19955803-A1.
XX
PD 23-MAY-2001.
XX
PF 19-NOV-1999; 99DE-1055803.
XX
PR 19-NOV-1999; 99DE-1055803.
XX
PA (KNOL) KNOLL AG.
XX
PI Herr D, Hahn A, Laux V;
XX
DR WPI; 2001-368371/39.
DR N-PSDB; AAH20940.
XX
PT Treatment or prevention of cardiac insufficiency and related
PT conditions, e.g. pulmonary congestion and dyspnoea, comprises
PT administration of heparanase inhibitor
XX
PS Disclosure; Page 11-13; 16pp; German.
XX
CC This invention describes a novel heparanase inhibitor which can be used
CC for the treatment or prevention of cardiac insufficiency and associated
CC indications, symptoms and/or malfunctions. The heparanase inhibitor of
CC the invention has cardiant, nephrotropic and hepatotropic activity. The
CC products of the invention can be used in human and veterinary medicine,
CC for the treatment or prevention of congestive heart failure e.g. primary
CC cardiomyopathy. Associated conditions treated or prevented with the
CC inhibitor are especially peripheral odemas, pulmonary and hepatic
CC congestion, dyspnoea, hydrothorax and ascites. Renal problems, e.g.
CC nocturia can also be treated. This sequence represents the human
CC heparanase protein described in the method of the invention.
XX
SQ Sequence 543 AA;

Query Match 99.9%; Score 2838; DB 22; Length 543;
Best Local Similarity 99.8%; Pred. No. 3.3e-273;
Matches 542; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy     541 ACI 543
           |||
Db     541 ACI 543
```

RESULT 11
AAY30124
ID AAY30124 standard; Protein; 588 AA.
XX
AC AAY30124;
XX
DT 14-OCT-1999 (first entry)
XX
DE A human protein with heparanase activity.
XX
KW Human; heparanase; heparan sulfate; trauma; autoimmune disease;
KW skin disease; cardiovascular disease; nervous system disease;
KW Alzheimer's disease; cancer; cancer metastasis; angiogenesis;
KW inflammation; arthritis.
XX
OS Homo sapiens.
XX
PN WO9940207-A1.
XX
PD 12-AUG-1999.
XX
PF 05-FEB-1999; 99WO-EP00777.
XX
PR 09-FEB-1998; 98GB-0002725.
XX
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Nakajima M, Toyoshima M;
XX
DR WPI; 1999-494300/41.
DR N-PSDB; AAX86671.
XX
PT New heparanase polypeptide useful for treating autoimmune diseases,
PT skin diseases, cardiovascular diseases and nervous system diseases
PT including Alzheimer's disease
XX
PS Claim 3; Page 29-31; 40pp; English.
XX
CC The present sequence represents a polypeptide with human heparanase
CC biological activity. Antagonists and inhibitors of the protein prevent

```
CC   it from degrading the extracellular matrix and releasing heparan sulfate
CC   from the extracellular matrix surface. The heparanase protein or the
CC   anti-heparanase antibody are used in pharmaceutical compositions for
CC   treating warm blooded animals suffering from a disease resulting from
CC   shortage or lack of the heparanase protein, or from excessive activity
CC   or over-expression of the heparanase protein, respectively. The
CC   heparanase protein is used in treating diseases such as trauma,
CC   autoimmune disease, skin diseases, cardiovascular diseases and nervous
CC   system diseases including Alzheimer's disease resulting from shortage or
CC   lack of polypeptide. The anti-heparanase antibody is used in treating
CC   the diseases like cancer, cancer metastasis, angiogenesis and
CC   inflammation including arthritis resulting from excessive activity or
CC   over expression of heparanase protein. The anti-heparanase antibody can
CC   be used to detect the presence or absence of polypeptide and its
CC   concentration.
XX
SQ   Sequence   588 AA;

  Query Match             99.9%;  Score 2838;  DB 20;  Length 588;
  Best Local Similarity   99.8%;  Pred. No. 3.8e-273;
  Matches  542;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 105

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 165

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 225

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     226 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 285

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     286 QLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 345

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     346 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 405

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     406 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 465

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 525

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     526 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 585

Qy     541 ACI 543
           |||
Db     586 ACI 588


RESULT 12
AAB88361
ID   AAB88361 standard; Protein; 543 AA.
XX
AC   AAB88361;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Human membrane or secretory protein clone PSEC0090.
XX
KW   Human; secretory protein; membrane protein; vaccine; gene therapy;
KW   rheumatoid arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   EP1067182-A2.
XX
PD   10-JAN-2001.
XX
PF   07-JUL-2000; 2000EP-0114090.
XX
PR   08-JUL-1999;   99JP-0194179.
PR   11-JAN-2000; 2000JP-0118775.
PR   02-MAY-2000; 2000JP-0183766.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Kawai Y,  Sugiyama T,  Hayashi K;
XX
DR   WPI; 2001-093989/11.
DR   N-PSDB; AAF93788.
XX
PT   Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT   gene therapy or as candidate target molecules in drug development -
XX
PS   Claim 1; SEQ ID 90; 609pp + CD ROM; English.
XX
CC   This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC   which encode human secretory or membrane proteins represented by
CC   AAB88317 - AAB88419. Included in the invention are primers
CC   AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
CC   cDNA sequences of the invention. The invention also includes methods for
CC   the production of antibodies directed against the proteins, and cDNA
CC   sequences, which can be used in vaccines. The polynucleotide sequences
CC   can be used in gene therapy. The polynucleotide sequences and the
CC   proteins they encode may be used in the prevention, treatment and
CC   diagnosis of diseases associated with inappropriate secretory
CC   protein/membrane protein expression. The nucleic acids and complementary
CC   sequences may also be used as DNA probes in diagnostic assays
CC   (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
CC   presence of similar nucleic acid sequences in samples. They may also be
CC   used to study the expression and function of secretory proteins/membrane
CC   polypeptides and their role in metabolism. The polypeptides may be used
CC   as antigens in the production of antibodies against them and in assays to
CC   identify modulators (agonists and antagonists) of expression and
CC   activity. The antibodies and antagonists may also be used as therapeutic
CC   agents to down regulate expression and activity. The antibodies may also
CC   be used as diagnostic agents for detecting the presence of the
CC   polypeptides in samples (e.g. by enzyme linked immunosorbant assay
CC   (ELISA). Examples of diseases which may be treated include rheumatoid
CC   arthritis and diabetes.
XX
SQ   Sequence   543 AA;

  Query Match             99.4%;  Score 2826;  DB 22;  Length 543;
  Best Local Similarity   99.4%;  Pred. No. 5.2e-272;
  Matches  540;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLDLIFGLNALLGTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
```

```
Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAHGGGAPLLSDTFA 360

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVA 540

Qy     541 ACI 543
           |||
Db     541 ACI 543
```

```
RESULT 13
AAY34173
ID   AAY34173 standard; Protein; 530 AA.
XX
AC   AAY34173;
XX
DT   15-NOV-1999  (first entry)
XX
DE   Human pre-proheparanase protein sequence.
XX
KW   Human; pre-proheparanase; platelet; wound healing; angiogenesis blocker;
KW   inflammation; psoriasis; diabetic retinopathy; solid tumour; arthritis;
KW   heparin degradation; anticoagulant neutralisation; asthma; CNS disease;
KW   inflammatory disease; vascular restenosis; atherosclerosis; diagnosis;
KW   tumour growth; fibroproliferative disorder; neurodegenerative disease;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9943830-A2.
XX
PD   02-SEP-1999.
XX
PF   18-FEB-1999;   99WO-US01489.
XX
PR   26-MAR-1998;   98US-0079401.
PR   24-FEB-1998;   98US-0075706.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Fairbanks MB,  Heinrikson RL,  Mildner AM;
XX
DR   WPI; 1999-540598/45.
DR   N-PSDB; AAZ11236.
XX
PT   New isolated platelet heparanase polypeptides, used to develop
PT   products for, e.g. wound healing and blocking angiogenesis
XX
PS   Claim 12; Fig 7; 57pp; English.
XX
CC   This sequence is the human pre-proheparanase of the invention. This
CC   sequence was isolated from human platelets. The heparanase can be used
CC   for identifying agents which alter heparanase activity. The heparanase
CC   can be used for wound healing or for blocking angiogenesis or
CC   inflammation. It can be used for treating e.g. psoriasis, diabetic
CC   retinopathy or solid tumours, or for the degradation of heparin and the
CC   neutralisation of heparin's anticoagulant properties during surgery.
CC   Inhibitors of heparanase activity can be used in the treatment of
CC   arthritis, asthma, and other inflammatory diseases, vascular restenosis,
CC   atherosclerosis, tumour growth and progression, fibroproliferative
CC   disorders, and central nervous system (CNS) and neurodegenerative
CC   diseases. The products can also be used for detection and diagnosis. The
CC   purified heparanase, both recombinantly produced human heparanase and
CC   heparanase isolated from human platelet activity, allows for the
CC   convenient selection of compounds having anti-heparanase activity,
CC   i.e. inhibitors of heparanase activity, by measuring inhibition of
CC   heparanase activity. Inhibition of heparanase activity can be measured by
CC   blocking heparanase-mediated release of radioactive fragments from in
CC   vivo radiolabelled (HSPG)/heparin.
XX
SQ   Sequence   530 AA;
```

```
  Query Match          97.3%;  Score 2764;  DB 20;  Length 530;
  Best Local Similarity   99.4%;  Pred. No. 7.4e-266;
  Matches  527;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      14 MLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFLI 73
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLLLGPLGPFSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFLI 60

Qy      74 LLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIP 133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSIP 120

Qy     134 PDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALLR 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 PDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALLR 180

Qy     194 TADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDYIQLHKLL 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db     181 TADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDFIQLHKLL 240

Qy     254 RKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLNP 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db     241 RKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATKEDFLNP 300

Qy     314 DVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSA 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 DVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLSA 360

Qy     374 RMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRV 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 RMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLRV 420

Qy     434 YLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSV 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 YLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKSV 480

Qy     494 QLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI 543
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 QLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI 530
```

```
RESULT 14
AAY17083
ID   AAY17083 standard; Protein; 532 AA.
XX
AC   AAY17083;
XX
DT   21-JUL-1999  (first entry)
XX
DE   Seq ID No: 15 of WO9921975.
XX
KW   Heparanase; endoglucuronidase; heparan sulfate proteoglycan; enzyme;
KW   metastasis; angiogenesis; wound healing; angioplasty-induced restenosis;
KW   arteriosclerosis; atherosclerosis; inflammation; tissue development;
KW   human; HSPG.
XX
OS   Homo sapiens.
XX
```

```
PN   WO9921975-A1.
XX
PD   06-MAY-1999.
XX
PF   28-OCT-1998;   98WO-AU00898.
XX
PR   09-DEC-1997;   97AU-0000812.
PR   28-OCT-1997;   97AU-0000062.
XX
PA   (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI   Freeman CG,  Hamdorf BJ,  Hulett MD,  Parish CR;
XX
DR   WPI; 1999-312956/26.
DR   N-PSDB; AAX37260.
XX
PT   Polynucleotides encoding mammalian endoglucuronidases, especially
PT   heparanases, useful to promote wound healing
XX
PS   Claim 6; Page 76-79; 112pp; English.
XX
CC   The invention relates to nucleic acid sequences that encode heparanase
CC   enzymes having  endoglucuronidase activity. Recombinant heparanases are
CC   capable of removing the HS side chain from heparan sulfate proteoglycan
CC   (HSPG). Sulfated oligosaccharides, sulphonates or HSPG can be used to
CC   inhibit heparanase, this is useful for treatment of a physiological or
CC   medical condition associated with elevated heparanase activity, such as
CC   metastasis, angiogenesis, wound healing, angioplasty-induced restenosis,
CC   arteriosclerosis, atherosclerosis and inflammation. The human, murine and
CC   rat heparanases can be used to enhance wound healing, especially
CC   associated with tissue development and repair. The conditions mentioned
CC   above can be diagnosed using specific antibodies, and also using primers
CC   and probes specific for the heparanase polynucleotides. Other uses of the
CC   heparanases include sequencing sulfated molecules such as HSPG.
XX
SQ   Sequence   532 AA;

  Query Match             96.3%;  Score 2737;  DB 20;  Length 532;
  Best Local Similarity   99.8%;  Pred. No. 3.6e-263;
  Matches  522;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MLLRSKPALPPPLMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVT 60

Qy      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IDANLATDPRFLILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQS 120

Qy     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QVNQDICKYGSIPPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCS 180

Qy     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLDLIFGLNALLRTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGS 240

Qy     241 QLGEDYIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QLGEDFIQLHKLLRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYL 300

Qy     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NGRTATREDFLNPDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFA 360

Qy     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGFMWLDKLGLSARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLM 420

Qy     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ASVQGSKRRKLRVYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLL 480

Qy     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLG 523
           |||||||||||||||||||||||||||||||||||||||||||||
Db     481 RPLGPHGLLSKSVQLNGLTLKMVDDQTLPPLMEKPLRPGSSLG 523


RESULT 15
ABB07815
ID   ABB07815 standard; protein; 527 AA.
XX
AC   ABB07815;
XX
DT   03-JUL-2002  (first entry)
XX
DE   Chicken signal peptide/human heparanase chimeric protein sequence.
XX
KW   Heparanase; catalytic; cytostatic; antiviral; antibacterial; enzyme;
KW   anti-protozoan; neuroprotective; heparin; chicken; human; chimeric.
XX
OS   Synthetic.
OS   Gallus gallus.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..19
FT                   /note= "chicken heparanase signal peptide"
FT   Protein         20..527
FT                   /note= "human heparanase mature protein"
XX
PN   US2002034810-A1.
XX
PD   21-MAR-2002.
XX
PF   16-AUG-2001; 2001US-0930218.
XX
PR   20-SEP-2000; 2000US-0666390.
XX
PA   (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX
PI   Goldshmidt O,  Pecker I,  Vlodavsky I,  Michal I,  Zcharia E;
XX
DR   WPI; 2002-338926/37.
DR   N-PSDB; ABL40753.
XX
PT   Nucleic acid encoding avian and reptile heparanase polypeptide is
PT   useful to treat various heparin-related disorders and the signal
PT   peptide is useful in production of membrane-targeted or secreted
PT   recombinant proteins
XX
PS   Disclosure; Page 26-28; 39pp; English.
XX
CC   The invention relates to an isolated avian and reptile nucleic acid,
CC   encoding a polypeptide with heparanase catalytic activity. The signal
CC   peptide of the nucleic acid can be used to express membrane-associated or
CC   secreted proteins in heterologous expression systems. The encoded
CC   polypeptides can be used to prevent tumour angiogenesis, metastasis and
CC   invasion, and to intervene with pathologies associated with impaired
CC   heparin-binding growth factors, cellular responses to heparin-binding
CC   growth factors and cytokines, cell interaction with plasma lipoproteins,
CC   cellular susceptibility to viral, protozoa and bacterial infections or
CC   disintegration of neurodegenerative plaques. The present sequence
CC   represents a chicken signal peptide/human heparanase chimeric protein
CC   sequence.
XX
SQ   Sequence   527 AA;

  Query Match             94.1%;  Score 2673.5;  DB 23;  Length 527;
  Best Local Similarity   96.8%;  Pred. No. 7.6e-257;
  Matches  514;  Conservative    4;  Mismatches    4;  Indels    9;  Gaps    1;

Qy      13 LMLLLLGPLGPLSPGALPRPAQAQDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFL 72
           |::|||         |:|    ||||||||||||||||||||||||||||||||||||
Db       6 LLVLLL---------AVPPRRTAQDVVDLDFFTQEPLHLVSPSFLSVTIDANLATDPRFL 56
```

```
Qy      73 ILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSI 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57 ILLGSPKLRTLARGLSPAYLRFGGTKTDFLIFDPKKESTFEERSYWQSQVNQDICKYGSI 116

Qy     133 PPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALL 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117 PPDVEEKLRLEWPYQEQLLLREHYQKKFKNSTYSRSSVDVLYTFANCSGLDLIFGLNALL 176

Qy     193 RTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDYIQLHKL 252
           ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db     177 RTADLQWNSSNAQLLLDYCSSKGYNISWELGNEPNSFLKKADIFINGSQLGEDFIQLHKL 236

Qy     253 LRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLN 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     237 LRKSTFKNAKLYGPDVGQPRRKTAKMLKSFLKAGGEVIDSVTWHHYYLNGRTATREDFLN 296

Qy     313 PDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLS 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     297 PDVLDIFISSVQKVFQVVESTRPGKKVWLGETSSAYGGGAPLLSDTFAAGFMWLDKLGLS 356

Qy     373 ARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLR 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     357 ARMGIEVVMRQVFFGAGNYHLVDENFDPLPDYWLSLLFKKLVGTKVLMASVQGSKRRKLR 416

Qy     433 VYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKS 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     417 VYLHCTNTDNPRYKEGDLTLYAINLHNVTKYLRLPYPFSNKQVDKYLLRPLGPHGLLSKS 476

Qy     493 VQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI 543
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     477 VQLNGLTLKMVDDQTLPPLMEKPLRPGSSLGLPAFSYSFFVIRNAKVAACI 527
```

Search completed: November 20, 2002, 11:36:11
Job time : 40 secs

=> dis his

(FILE 'HOME' ENTERED AT 13:53:38 ON 21 NOV 2002)

FILE 'MEDLINE, CAPLUS, EMBASE, BIOSIS' ENTERED AT 13:54:35 ON 21 NOV 2002

L1 406 S PECKER I?/AU OR VLODASKY I?/AU OR FRIEDMAN Y?/AU OR PERETS T?
L2 60 S L1 AND HEPARANASE
L3 12 S L1 AND (HEPARANASE (10N) ANTIBOD?)
L4 9 DUP REM L3 (3 DUPLICATES REMOVED)
L5 79 S (HEPARANASE (10N) ANTIBOD?)
L6 70 S L5 NOT L4
L7 18 S L6 AND PD<19970902
L8 7 DUP REM L7 (11 DUPLICATES REMOVED)

WEST

Create A Case

Select?	Database	Query	Plural	Op	Thesaurus	Set Name
☑	USPT	(5968822)[PN]	YES	OR		L1
☑	USPT,PGPB,JPAB,EPAB,DWPI	(pecker)[in[] or (vlodasky)[in] or (friedman)[in] or (perets)[in]	YES	OR		L2
☑	USPT,PGPB,JPAB,EPAB,DWPI	L2 and heparanase	YES	OR		L3
☑	USPT,PGPB,JPAB,EPAB,DWPI	antibod$4and heparanase	YES	OR		L4
☑	USPT,PGPB,JPAB,EPAB,DWPI	L4 and @ad<19970902	YES	OR		L5
☑	USPT,PGPB,JPAB,EPAB,DWPI	antibod$4 near heparanase	YES	OR		L6
☑	USPT,PGPB,JPAB,EPAB,DWPI	L6 and @ad<19970902	YES	OR		L7

Please enter the case name: 09759207

Clear All | Reset | Create Case | Cancel

Help | Main Menu | Logout

Rules for naming Cases

- Case names can only contain alphanumeric characters including underscore (_).
- Any other special characters or punctuation characters will be automatically removed prior to saving the case.
- All white space characters will be replaced by an underscore.

Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:ssspta1644axd

PASSWORD:
TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * * * * * * Welcome to STN International * * * * * * * * * *

```
 NEWS  1           Web Page URLs for STN Seminar Schedule - N. America
 NEWS  2  Apr 08   "Ask CAS" for self-help around the clock
 NEWS  3  Apr 09   BEILSTEIN: Reload and Implementation of a New Subject Area
 NEWS  4  Apr 09   ZDB will be removed from STN
 NEWS  5  Apr 19   US Patent Applications available in IFICDB, IFIPAT, and IFIUDB
 NEWS  6  Apr 22   Records from IP.com available in CAPLUS, HCAPLUS, and ZCAPLUS
 NEWS  7  Apr 22   BIOSIS Gene Names now available in TOXCENTER
 NEWS  8  Apr 22   Federal Research in Progress (FEDRIP) now available
 NEWS  9  Jun 03   New e-mail delivery for search results now available
 NEWS 10  Jun 10   MEDLINE Reload
 NEWS 11  Jun 10   PCTFULL has been reloaded
 NEWS 12  Jul 02   FOREGE no longer contains STANDARDS file segment
 NEWS 13  Jul 22   USAN to be reloaded July 28, 2002;
                   saved answer sets no longer valid
 NEWS 14  Jul 29   Enhanced polymer searching in REGISTRY
 NEWS 15  Jul 30   NETFIRST to be removed from STN
 NEWS 16  Aug 08   CANCERLIT reload
 NEWS 17  Aug 08   PHARMAMarketLetter(PHARMAML) - new on STN
 NEWS 18  Aug 08   NTIS has been reloaded and enhanced
 NEWS 19  Aug 19   Aquatic Toxicity Information Retrieval (AQUIRE)
                   now available on STN
 NEWS 20  Aug 19   IFIPAT, IFICDB, and IFIUDB have been reloaded
 NEWS 21  Aug 19   The MEDLINE file segment of TOXCENTER has been reloaded
 NEWS 22  Aug 26   Sequence searching in REGISTRY enhanced
 NEWS 23  Sep 03   JAPIO has been reloaded and enhanced
 NEWS 24  Sep 16   Experimental properties added to the REGISTRY file
 NEWS 25  Sep 16   Indexing added to some pre-1967 records in CA/CAPLUS
 NEWS 26  Sep 16   CA Section Thesaurus available in CAPLUS and CA
 NEWS 27  Oct 01   CASREACT Enriched with Reactions from 1907 to 1985
 NEWS 28  Oct 21   EVENTLINE has been reloaded
 NEWS 29  Oct 24   BEILSTEIN adds new search fields
 NEWS 30  Oct 24   Nutraceuticals International (NUTRACEUT) now available on STN
 NEWS 31  Oct 25   MEDLINE SDI run of October 8, 2002
 NEWS 32  Nov 18   DKILIT has been renamed APOLLIT

 NEWS EXPRESS  October 14 CURRENT WINDOWS VERSION IS V6.01,
               CURRENT MACINTOSH VERSION IS V6.0a(ENG) AND V6.0Ja(JP),
               AND CURRENT DISCOVER FILE IS DATED 01 OCTOBER 2002
 NEWS HOURS    STN Operating Hours Plus Help Desk Availability
 NEWS INTER    General Internet Information
 NEWS LOGIN    Welcome Banner and News Items
 NEWS PHONE    Direct Dial and Telecommunication Network Access to STN
 NEWS WWW      CAS World Wide Web Site (general information)
```

Enter NEWS followed by the item number or name to see news on that specific topic.

research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

* * * * * * * * * * * * * * * STN Columbus * * * * * * * * * * * * * * * *

FILE 'HOME' ENTERED AT 13:53:38 ON 21 NOV 2002

=> file medline caplus embase biosis
COST IN U.S. DOLLARS SINCE FILE ENTRY TOTAL SESSION
FULL ESTIMATED COST 0.42 0.42

FILE 'MEDLINE' ENTERED AT 13:54:35 ON 21 NOV 2002

FILE 'CAPLUS' ENTERED AT 13:54:35 ON 21 NOV 2002
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'EMBASE' ENTERED AT 13:54:35 ON 21 NOV 2002
COPYRIGHT (C) 2002 Elsevier Science B.V. All rights reserved.

FILE 'BIOSIS' ENTERED AT 13:54:35 ON 21 NOV 2002
COPYRIGHT (C) 2002 BIOLOGICAL ABSTRACTS INC.(R)

=> s pecker I?/au or vlodasky I?/au or friedman Y?/au or Perets T?/au
L1 406 PECKER I?/AU OR VLODASKY I?/AU OR FRIEDMAN Y?/AU OR PERETS T?/AU

=> s l1 and heparanase
L2 60 L1 AND HEPARANASE

=> s l1 and (heparanase (10N) antibod?
MISSING OPERATOR 'EPARANASE (10N'
The search profile that was entered contains terms or nested terms that are not separated by a logical operator.

=> s l1 and (heparanase (10N) antibod?)
MISSING OPERATOR 'EPARANASE (10N'
The search profile that was entered contains terms or nested terms that are not separated by a logical operator.

=> s l1 and (heparanase (10N) antibod?)
MISSING OPERATOR 'EPARANASE (10N'
The search profile that was entered contains terms or nested terms that are not separated by a logical operator.

=> s (heparanase (10N) antibod?)
MISSING OPERATOR 'EPARANASE (10N'
The search profile that was entered contains terms or nested terms that are not separated by a logical operator.

=> s (heparanase (10N) antibod?)
MISSING OPERATOR 'EPARANASE (10N'
The search profile that was entered contains terms or nested terms that are not separated by a logical operator.

=> s l1 and (heparanase (10N) antibod?)
L3 12 L1 AND (HEPARANASE (10N) ANTIBOD?)

=> dup rem l3

PROCESSING COMPLETED FOR L3
L4 9 DUP REM L3 (3 DUPLICATES REMOVED)

=> dis l4 1-9 ibib abs

L4 ANSWER 1 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2002:429402 CAPLUS
DOCUMENT NUMBER: 137:5014
TITLE: Monoclonal anti-**heparanase** **antibodies** for treating conditions related to heparan sulfate proteoglycan-associated growth factor, chemokine, cytokine or degradative enzyme
INVENTOR(S): Peretz, Tuvia; Miron, Daphna; Shlomi, Yinon; **Pecker, Iris**; Ayal-Hershkovitz, Maty; Vlodavsky, Isreal; Friedmann, Yael
PATENT ASSIGNEE(S): Israel
SOURCE: U.S. Pat. Appl. Publ., 11 pp., Cont.-in-part of U.S. Ser. No. 71,618, abandoned.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 15
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2002068061	A1	20020606	US 1998-186200	19981104
US 5968822	A	19991019	US 1997-922170	19970902
WO 2000025817	A1	20000511	WO 1999-US25451	19991028
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
EP 1126878	A1	20010829	EP 1999-956781	19991028
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
AU 751170	B2	20020808	AU 2000-13314	19991028
NO 2001002190	A	20010612	NO 2001-2190	20010503

PRIORITY APPLN. INFO.:
US 1997-922170 A2 19970902
US 1998-71618 B2 19980501
US 1998-186200 A 19981104
WO 1999-US25451 W 19991028

AB Monoclonal **antibodies**, neutralizing **antibodies**, humanized **antibodies** specific to **heparanase** protein or an immunogenic portion thereof are disclosed. These **antibodies** are useful for inhibiting **heparanase** activity and for treating conditions assocd. with altered function of a HSPG-assocd. biol. effector mol., e.g. growth factor, chemokine, cytokine, or degradative enzyme. The condition is angiogenesis, cell proliferation, tumor cell proliferation, invasion of circulating tumor cell, metastasis, inflammatory disorders and autoimmune diseases.

L4 ANSWER 2 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2002:131509 CAPLUS
DOCUMENT NUMBER: 136:195300
TITLE: Genetically modified cells and methods for expressing recombinant human heparanase and methods of its purification
INVENTOR(S): Ayal-Hershkovitz, Maty; Moskowitz, Haim; Miron,

Daphna; Gilboa, Ayelet; Mimon, Madelene; Ben-Artzi, Hanna; Yacoby-Zeevi, Oron; **Pecker, Iris**; Peleg, Yoav; Schlomi, Yinon

PATENT ASSIGNEE(S): Insight Strategy & Marketing Ltd., Israel

SOURCE: U.S., 66 pp., Cont.-in-part of U.S. Ser. No. 71,618, abandoned.
CODEN: USXXAM

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 15

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6348344	B1	20020219	US 1999-260038	19990302
US 5968822	A	19991019	US 1997-922170	19970902
US 6177545	B1	20010123	US 1998-71739	19980501
CA 2329142	AA	19991111	CA 1999-2329142	19990429
WO 9957244	A1	19991111	WO 1999-US9256	19990429
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
AU 9937705	A1	19991123	AU 1999-37705	19990429
EP 1076689	A1	20010221	EP 1999-920135	19990429
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
JP 2002513560	T2	20020514	JP 2000-547200	19990429
US 6475763	B1	20021105	US 2000-487716	20000119
US 6426209	B1	20020730	US 2000-635923	20000810
NO 2000005100	A	20001228	NO 2000-5100	20001010

PRIORITY APPLN. INFO.:

US 1997-922170	A2	19970902
US 1998-71618	B2	19980501
US 1998-71739	A2	19980501
US 1999-260038	A	19990302
WO 1999-US9256	W	19990429
US 2000-487716	A1	20000119

AB Bacterial, yeast, and animal cells and methods for overexpressing recombinant heparanase in cellular systems, methods of purifying recombinant heparanase therefrom and modified heparanase species which serve as precursors for generating highly active heparanase by proteolysis are provided. Thus, cloning of human heparanase cDNA into baculovirus-infected High 5 and Sf21 cells yielded 0.44 and 0.16 mg enzyme/mL, resp. Enzyme purifn. is achieved by cation-exchange chromatog. on Source-S or affinity chromatog. with anti-native **heparanase antibodies**. Highly active partially proteolytically cleaved forms of heparanase were identified. This led to the construction of recombinant heparanase contg. (1) an enterokinase cleavage site (Ser-Gln-Val-Asn-Gln) leading to cleavage between residues 119 and 120, or (2) a cathepsin L cleavage site leading to cleavage between residues 157 and 158.

REFERENCE COUNT: 24 THERE ARE 24 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 3 OF 9 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2001:12473 CAPLUS

DOCUMENT NUMBER: 134:96257

TITLE: Protein and cDNA sequences of a novel human heparanase gene hnhp1 and its splicing variants

INVENTOR(S): **Pecker, Iris**; Michal, Israel; Itzhaki, Hanan
PATENT ASSIGNEE(S): Insight Strategy & Marketing Ltd., Israel
SOURCE: PCT Int. Appl., 67 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001000643	A2	20010104	WO 2000-IL358	20000619
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
EP 1212341	A1	20020612	EP 2000-937164	20000619
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL				
NO 2001005526	A	20011218	NO 2001-5526	20011112

PRIORITY APPLN. INFO.: US 1999-140801P P 19990625
WO 2000-IL358 W 20000619

AB The invention provides protein and cDNA sequences of a novel human heparanase gene hnhp1 and two variants resulted from alternative splicing. The longest clone is 2060 nucleotide long and it contains an open reading frame of 1776 nucleotides, which encodes a polypeptide of 592 amino acids, with a calcd. mol. wt. of 66.5 kDa. The two shorter forms contain an in frame deletion as a result of alternative splicing, one is 162 nucleotides (nt473-634) corresponding to amino acids 150-203, and one is 336 nucleotides (nt473-808) corresponding to amino acids 150-261. The hnhp1 gene is mapped to chromosome 10, next to the marker SHGC-57721. The tissue distribution of hnhp1 transcripts is detd. The invention also relates to constructing hnhp1 gene expression vector to produce recombinant proteins in mammalian cells, which may have **heparanase** or other glycosyl hydrolase activity, its **antibodies**, and antisense oligonucleotide and ribozymes for gene modulation and therapeutic use.

REFERENCE COUNT: 52 THERE ARE 52 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 4 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:57239 CAPLUS
DOCUMENT NUMBER: 134:128217
TITLE: Heparanase specific molecular probes and their use in research and medical applications
INVENTOR(S): **Pecker, Iris**; Vlodavsky, Israel; **Friedman, Yael**; **Perets, Tuvia**
PATENT ASSIGNEE(S): Insight Strategy & Marketing Ltd., Israel
SOURCE: U.S., 41 pp., Cont.-in-part of U.S. 5,968,822.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 15
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6177545	B1	20010123	US 1998-71739	19980501
US 5968822	A	19991019	US 1997-922170	19970902

US 6348344	B1	20020219	US 1999-260038	19990302	
WO 9957153	A1	19991111	WO 1999-US9255	19990429	

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

AU 9938706	A1	19991123	AU 1999-38706	19990429
EP 1073682	A1	20010207	EP 1999-921513	19990429

R: AT, BE, CH, DE, DK, ES, FR, GB, IT, LI, NL, SE, IE, FI

JP 2002512533	T2	20020423	JP 1999-555528	19990429
US 2002114801	A1	20020822	US 1999-322977	19990601
NO 9906229	A	20000224	NO 1999-6229	19991215
US 6475763	B1	20021105	US 2000-487716	20000119
US 6426209	B1	20020730	US 2000-635923	20000810
US 2002004585	A1	20020110	US 2001-759207	20010116
US 2002102619	A1	20020801	US 2001-944602	20010904

PRIORITY APPLN. INFO.:

US 1997-922170	A2	19970902
US 1998-71618	B2	19980501
US 1998-71739	A2	19980501
US 1999-260038	A1	19990302
WO 1999-US9255	W	19990429
US 1999-322977	A1	19990601
US 2000-487716	A1	20000119
US 2001-759207	A1	20010116

AB A variety of heparanase specific mol. probes which can be used for research and medical applications including diagnosis and therapy. Specific applications include the use of a heparanase specific mol. probe for detection of the presence, absence or level of heparanase expression; the use of a heparanase specific mol. probe for therapy of a condition assocd. with expression of heparanase; the use of a heparanase specific mol. probe for quantification of heparanase in a body fluid; the use of a heparanase specific mol. probe for targeted drug delivery; and the use of a heparanase specific mol. probe as a therapeutic agent.

REFERENCE COUNT: 23 THERE ARE 23 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 5 OF 9 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2000:314574 CAPLUS

DOCUMENT NUMBER: 132:333392

TITLE: **Heparanase** activity neutralizing anti-**heparanase** monoclonal **antibody**

INVENTOR(S): Peretz, Tuvia; Miron, Daphna; Shlomi, Yinon; **Pecker, Iris**; Ayal-Hershkovitz, Maty; **Friedman, Yael**; Vlodavsky, Israel

PATENT ASSIGNEE(S): Insight Strategy & Marketing Ltd., Israel; Hadasit Medical Research Services & Development Ltd.; Friedman, Mark M.

SOURCE: PCT Int. Appl., 28 pp.
CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 15

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000025817	A1	20000511	WO 1999-US25451	19991028

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS,

JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

US 2002068061	A1	20020606	US 1998-186200	19981104
EP 1126878	A1	20010829	EP 1999-956781	19991028
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
AU 751170	B2	20020808	AU 2000-13314	19991028
NO 2001002190	A	20010612	NO 2001-2190	20010503

PRIORITY APPLN. INFO.:

US 1998-186200	A	19981104
US 1997-922170	A2	19970902
US 1998-71618	B2	19980501
WO 1999-US25451	W	19991028

AB A monoclonal **antibody** elicited by a **heparanase** protein or an immunogenic portion thereof, the monoclonal **antibody** specifically inhibits **heparanase** activity. The **heparanase**-specific monoclonal **antibody** may be human or humanized **antibody** and is useful for treating conditions assocd. with altered function of a heparan sulfate proteoglycan-assocd. biol. effector mol. such as growth factor, chemokine, cytokine and degradative enzyme. The condition is selected from the group consisting of angiogenesis, cell proliferation, tumor, metastasis, inflammatory disorders and autoimmune conditions.

REFERENCE COUNT: 6 THERE ARE 6 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 6 OF 9 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2000:53938 CAPLUS
DOCUMENT NUMBER: 132:102821
TITLE: Method of screening for potential anti-metastatic and anti-inflammatory agents using mammalian heparanase as a probe
INVENTOR(S): Ben-Artzi, Hanna; Ayal-Hershkovitz, Maty; Vlodavsky, Israel; **Pecker, Iris**; Peleg, Yoav; Miron, Daphna
PATENT ASSIGNEE(S): Insight Strategy & Marketing Ltd., Israel; Hadasit Medical Research Services & Development Ltd.; Friedman, Mark M.
SOURCE: PCT Int. Appl., 70 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 15
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000003036	A1	20000120	WO 1999-US15643	19990712
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 6190875	B1	20010220	US 1998-113168	19980710
CA 2335382	AA	20000120	CA 1999-2335382	19990712
AU 9948697	A1	20000201	AU 1999-48697	19990712

EP 1097241 A1 20010509 EP 1999-932382 19990712
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO
JP 2002520029 T2 20020709 JP 2000-559256 19990712
NO 2001000136 A 20010309 NO 2001-136 20010109
PRIORITY APPLN. INFO.: US 1998-113168 A 19980710
US 1997-922170 A2 19970902
US 1998-109386 B2 19980702
WO 1999-US15643 W 19990712

AB Qual. and quant. methods are provided for testing an agent for its potential at inhibiting glycosidase catalytic activity, the methods including interacting a glycosidase enzyme with a glycosidase substrate in a presence of the agent and qual. or quant. evaluating an effect of the agent on the catalytic activity of the glycosidase enzyme toward the glycosidase substrate. Preferably the glycosidase enzyme is a heparanase enzyme and the glycosidase substrate is, resp., a heparanase substrate.

REFERENCE COUNT: 7 THERE ARE 7 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 7 OF 9 MEDLINE DUPLICATE 1
ACCESSION NUMBER: 2001009022 MEDLINE
DOCUMENT NUMBER: 20476203 PubMed ID: 11021821
TITLE: Expression of heparanase in normal, dysplastic, and neoplastic human colonic mucosa and stroma. Evidence for its role in colonic tumorigenesis.
AUTHOR: Friedmann Y; Vlodavsky I; Aingorn H; Aviv A; Peretz T; **Pecker I**; Pappo O
CORPORATE SOURCE: Departments of Oncology and Pathology, Hadassah-Hebrew University Hospital, Jerusalem, and InSight Ltd., Rabin Science Park, Rehovot, Israel.
SOURCE: AMERICAN JOURNAL OF PATHOLOGY, (2000 Oct) 157 (4) 1167-75. Journal code: 0370502. ISSN: 0002-9440.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
ENTRY MONTH: 200010
ENTRY DATE: Entered STN: 20010322
Last Updated on STN: 20010322
Entered Medline: 20001025

AB The human heparanase gene, an endo-beta-glucuronidase that cleaves heparan sulfate at specific intrachain sites, has recently been cloned and shown to function in tumor progression and metastatic spread. Antisense digoxigenin-labeled **heparanase** RNA probe and monoclonal anti-human **heparanase antibodies** were used to examine the expression of the **heparanase** gene and protein in normal, dysplastic, and neoplastic human colonic mucosa. To our knowledge, this is the first systematic study of heparanase expression in human colon cancer. Both the heparanase gene and protein were expressed at early stages of neoplasia, already at the stage of adenoma, but were practically not detected in the adjacent normal-looking colon epithelium. Gradually increasing expression of heparanase was evident as the cells progressed from severe dysplasia through well-differentiated to poorly differentiated colon carcinoma. Deeply invading colon carcinoma cells showed the highest levels of the heparanase mRNA and protein associated with expression of both the gene and enzyme by adjacent desmoplastic stromal fibroblasts. A high expression was also found in colon carcinoma metastases to lung, liver, and lymph nodes, as well as in the accompanying stromal fibroblasts. Moreover, extracts derived from tumor tissue expressed much higher levels of the heparanase protein and activity as compared to the normal colon tissue. In all specimens, the heparanase gene and protein exhibited the same pattern of expression. These results suggest a role of heparanase in colon cancer progression and may have both prognostic and therapeutic applications.

L4 ANSWER 8 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1999:723147 CAPLUS
DOCUMENT NUMBER: 131:332967
TITLE: Genetically modified cells and methods for expressing recombinant heparanase and methods of purifying same
INVENTOR(S): Ben-Artzi, Hanna; Ayal-Hershkovitz, Maty; Yacoby-Zeevi, Oron; **Pecker, Iris**; Peleg, Yoav; Shlomi, Yinon
PATENT ASSIGNEE(S): Insight Strategy & Marketing Ltd., Israel; Friedman, Mark, M.
SOURCE: PCT Int. Appl., 118 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 15
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9957244	A1	19991111	WO 1999-US9256	19990429
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 6348344	B1	20020219	US 1999-260038	19990302
CA 2329142	AA	19991111	CA 1999-2329142	19990429
AU 9937705	A1	19991123	AU 1999-37705	19990429
EP 1076689	A1	20010221	EP 1999-920135	19990429
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
JP 2002513560	T2	20020514	JP 2000-547200	19990429
NO 2000005100	A	20001228	NO 2000-5100	20001010

PRIORITY APPLN. INFO.:
US 1998-71618 A 19980501
US 1999-260038 A 19990302
US 1997-922170 A2 19970902
US 1998-71739 A2 19980501
WO 1999-US9256 W 19990429

AB Bacterial, yeast and animal cells and methods for overexpressing recombinant heparanase in cellular systems, methods of purifying recombinant heparanase therefrom and modified heparanase species which serve as precursors for generating highly active heparanase by proteolysis. Heparanase is a glycosylated enzyme involved in catabolism of certain glycosaminoglycans, in tumor cell invasion and metastasis, and possibly in angiogenesis. It has potential therapeutic applications for viral infection, neurodegenerative diseases, restenosis, and atherosclerosis. A signal peptide was incorporated for effective protein secretion in yeast and bacteria and insect and mammalian cells. Protein secretion is achieved by induction by thrombin and calcium ionophores and immune complexes and antigens and mitogens. This work describes prodn. of heparanase on a biotechnol. scale of at least half a liter growth medium by affinity purifn. This large scale propagation of animal cells is described in a Spinner-basket bioreactor. The heparanase enzyme is activated by digestion with a protease such as cathepsin L or trypsin at appropriate pH. A correctly folded catalytically active heparanase is generated.

REFERENCE COUNT: 2 THERE ARE 2 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 9 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1999:723067 CAPLUS
DOCUMENT NUMBER: 131:350261
TITLE: Heparanase specific molecular probes and their use in research and medical applications
INVENTOR(S): **Pecker, Iris**; Vlodavsky, Israel; **Friedman, Yael**; **Perets, Tuvia**
PATENT ASSIGNEE(S): Insight Strategy & Marketing Ltd., Israel; Hadasit Medical Research Services & Development Ltd.; Friedman, Mark, M.
SOURCE: PCT Int. Appl., 90 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 15
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9957153	A1	19991111	WO 1999-US9255	19990429
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 6177545	B1	20010123	US 1998-71739	19980501
AU 9938706	A1	19991123	AU 1999-38706	19990429
EP 1073682	A1	20010207	EP 1999-921513	19990429
R: AT, BE, CH, DE, DK, ES, FR, GB, IT, LI, NL, SE, IE, FI				
JP 2002512533	T2	20020423	JP 1999-555528	19990429
NO 9906229	A	20000224	NO 1999-6229	19991215

PRIORITY APPLN. INFO.:
US 1998-71739 A 19980501
US 1997-922170 A2 19970902
WO 1999-US9255 W 19990429

AB A variety of heparanase specific mol. probes which can be used for research and medical applications including diagnosis and therapy. Specific applications include the use of a heparanase specific mol. probe for detection of the presence, absence or level of heparanase expression; the use of a heparanase specific mol. probe for therapy of a condition assocd. with expression of heparanase; the use of a heparanase specific mol. probe for quantification of heparanase in a body fluid; the use of a heparanase specific mol. probe for targeted drug delivery; and the use of a heparanase specific mol. probe as a therapeutic agent.

REFERENCE COUNT: 14 THERE ARE 14 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

```
=> s (heparanase (10N) antibod?)
L5          79 (HEPARANASE (10N) ANTIBOD?)

=> s l5 not l4
L6          70 L5 NOT L4

=> s l6 and PD<19970902
'19970902' NOT A VALID FIELD CODE
   3 FILES SEARCHED...
L7          18 L6 AND PD<19970902

=> dup rem l7
PROCESSING COMPLETED FOR L7
```

L8 7 DUP REM L7 (11 DUPLICATES REMOVED)

=> dis l8 1-7 ibib abs

L8 ANSWER 1 OF 7 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 1
ACCESSION NUMBER: 1997:763245 CAPLUS
DOCUMENT NUMBER: 128:47265
TITLE: Major colocalization of the extracellular-matrix degradative enzymes heparanase and gelatinase in tertiary granules of human neutrophils
AUTHOR(S): Mollinedo, Faustino; Nakajima, Motowo; Llorens, Ana; Barbosa, Enrique; Callejo, Sagrario; Gajate, Consuelo; Fabra, Angels
CORPORATE SOURCE: Facultad de Medicina, Laboratory of Signal Transduction and Leucocyte Biology, Instituto de Biologia y Genetica Molecular, Consejo Superior de Investigaciones Cientificas-Universidad de Valladolid, Valladolid, E-47005, Spain
SOURCE: Biochemical Journal (**1997**), 327(3), 917-923
CODEN: BIJOAK; ISSN: 0264-6021
PUBLISHER: Portland Press Ltd.
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The expression of cell-surface adhesion proteins and the release of extracellular-matrix degradative enzymes constitute crucial processes for the attachment of neutrophils to the endothelium and for the subsequent extravasation of these cells through the endothelial layer. We have analyzed in resting human neutrophils the subcellular localization of heparanase, a heparan-sulfate-degrading endoglycosidase that can degrade basement-membrane components, thereby facilitating neutrophil passage into the tissue during an inflammatory reaction. By subcellular fractionation of postnuclear supernatants from resting human neutrophils on continuous gradients, we have found that heparanase activity was mainly located in gelatinase-contg. tertiary granules. Using a specific **antibody**, the 96-kDa **heparanase** protein was further located in the gelatinase-rich subcellular fractions. Following immunoblotting and immunopptn. anal. in the distinct subcellular fractions, we also found colocalization of heparanase and Mol (CD11b/CD18), a leukocyte integrin involved in the attachment of neutrophils to the endothelium, in the fractions enriched in gelatinase-contg. tertiary granules. Treatment of human neutrophils with tumor necrosis factor or granulocyte/macrophage colony-stimulating factor induced an increase in the CD11b/CD18 cell-surface expression, as well as the release of both gelatinase (matrix metalloproteinase-9) and heparanase, but not of other granule markers, indicating a major co-localization of gelatinase, heparanase and CD11b/CD18 in the same organelle. Furthermore, confocal laser scanning microscopy using specific **antibodies** against gelatinase and **heparanase** revealed a major co-localization of both enzymes in intracellular cytoplasmic granules. The major localization of heparanase and CD11b/CD18 in the gelatinase-contg. tertiary granule supports the notion that mobilization of this organelle can regulate extravasation on human neutrophils.

L8 ANSWER 2 OF 7 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.DUPLICATE 2
ACCESSION NUMBER: 97266746 EMBASE
DOCUMENT NUMBER: 1997266746
TITLE: Subendothelial retention of lipoprotein (a). Evidence that reduced heparan sulfate promotes lipoprotein binding to subendothelial matrix.
AUTHOR: Pillarisetti S.; Paka L.; Obunike J.C.; Berglund L.; Goldberg I.J.
CORPORATE SOURCE: Dr. S. Pillarisetti, Department of Medicine, Columbia University, College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032, United States.

ps42@columbia.edu
SOURCE: Journal of Clinical Investigation, (1997) 100/4 (867-874).
Refs: 59
ISSN: 0021-9738 CODEN: JCINAO
COUNTRY: United States
DOCUMENT TYPE: Journal; Article
FILE SEGMENT: 005 General Pathology and Pathological Anatomy
LANGUAGE: English
SUMMARY LANGUAGE: English

AB Vessel wall subendothelial extracellular matrix, a dense mesh formed of collagens, fibronectin, laminin, and proteoglycans, has important roles in lipid and lipoprotein retention and cell adhesion. In atherosclerosis, vessel wall heparan sulfate proteoglycans (HSPG) are decreased and we therefore tested whether selective loss of HSPG affects lipoprotein retention. A matrix synthesized by aortic endothelial cells and a commercially available matrix (Matrigel; Becton Dickinson Inc., Rutherford, NJ) were used. Treatment of matrix with heparinase/heparitinase (1 U/ml each) increased LDL binding by .apprx. 1.5-fold. Binding of lipoprotein (a) [Lp(a)] to both subendothelial matrix and Matrigel.RTM. increased 2-10-fold when the HSPG were removed by heparinase treatment. Incubation of endothelial cells with oxidized LDL (OxLDL) or lysolecithin resulted in decreased matrix proteoglycans and increased Lp(a) retention by matrix. The effect of OxLDL or lysolecithin on endothelial PG was abolished in the presence of HDL. The decrease in matrix HSPG was associated with production of a heparanase-like activity by OxLDL-stimulated endothelial cells. To test whether removal of HSPG exposes fibronectin, a candidate Lp(a) binding protein in the matrix, antifibronectin antibodies were used. The increased Lp(a) binding after HSPG removal was inhibited 60% by antifibronectin antibodies. Similarly, the increased Lp(a) binding to matrix from OxLDL-treated endothelial cells was inhibited by antifibronectin **antibodies**. We hypothesize that atherogenic lipoproteins stimulate endothelial cell production of **heparanase**. This enzyme reduces HSPG which in turn promotes Lp(a) retention.

L8 ANSWER 3 OF 7 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 3
ACCESSION NUMBER: 1997:194122 CAPLUS
DOCUMENT NUMBER: 126:262494
TITLE: Human prostate carcinoma cells produce extracellular heparanase
AUTHOR(S): Kosir, Mary Ann; Quinn, Christiane C. V.; Zukowski, Kim L.; Grignon, David J.; Ledbetter, Steven
CORPORATE SOURCE: VA Medical Center, Surgical Service, Detroit, MI, 48201, USA
SOURCE: Journal of Surgical Research (1997), 67(1), 98-105
CODEN: JSGRA2; ISSN: 0022-4804
PUBLISHER: Academic
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The degrdn. of heparan sulfate proteoglycan (HSPG) in basement membranes (BM) has been previously suggested to be accomplished by an endoglycosidase activity called heparanase which has not been isolated outside of platelets. HSPG degrdn. by heparanase has been assocd. with tumor cell invasion, angiogenesis, and growth factor function. In this study, we identify heparanase activity biochem. and immunol. in malignant human prostate carcinoma cells (PC-3M), linking platelet heparanase probes with the tumor heparanase activity obsd. Concd. conditioned medium from PC-3M cells was analyzed by a heparin-Sepharose affinity column. Three peaks eluted with 0.15, 0.35, and 0.5 M NaCl. Each peak was analyzed by incubation with 3H-labeled heparin as well as [3H]HSPG from EHS tumor BM. The 0.5 M peak material degraded [3H]-heparin by 17.2%, with little addnl. degrdn. by the other peaks in comparison to the conditioned medium from

which they were obtained. Likewise, the same amt. of the 0.5 M peak accounted for the majority of degrdn. (30.8%) of 3H-labeled HSPG. Interestingly, for the same amt. of 0.5 M peak material, significantly more HSPG was degraded than heparin under the same conditions. In addn., carrageenan-.lambda., an inhibitor of glycanase, completely inhibited the degrdn. of heparin and heparan sulfate proteoglycan by the 0.5 M peak. Using **antibody** to the N-terminus domain of platelet **heparanase**, a 60-kDa protein was identified by immunoblot in 0.5 M peak material. Addnl., immunohistochem. staining of human prostate carcinoma specimens showed granular staining at or near the cell membrane and near the luminal surface using **antibody** to the N-terminus and C-terminus domains of platelet **heparanase**. In summary, human prostate carcinoma cells show heparanase activity in conditioned medium that degrades heparin and BM HSPG and is detected by **antibody** to platelet **heparanase**. In addn., the membrane-assocd. staining in tissue sections of prostate cancer strongly correlates with the biochem. and immunol. detection in conditioned medium of human PC-3M cells.

L8 ANSWER 4 OF 7 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 4
ACCESSION NUMBER: 1994:627556 CAPLUS
DOCUMENT NUMBER: 121:227556
TITLE: immunoselection of GRP94/endoplasmin from a KNRK cell-specific .lambda.gt11 library using **antibodies** directed against a putative **heparanase** amino-terminal peptide
AUTHOR(S): De Vouge, Michael W.; Yamazaki, Amy; Bennett, Steffany A.L.; Chen, Jia Hua; Shwed, Philip S.; Couture, Chantal; Birnboim, H. Chaim
CORPORATE SOURCE: Ottawa Reg. Cancer Cent., Ottawa, ON, K1H 8L6, Can.
SOURCE: International Journal of Cancer (**1994**), 56(2), 286-94
CODEN: IJCNAW; ISSN: 0020-7136
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Induction of an invasive phenotype by metastatic tumor cells results in part from inappropriate expression of extracellular matrix-degrading enzymes normally involved in embryonic morphogenesis, tissue remodelling, angiogenesis and wound healing. Such enzymes include endoglycosidases that degrade heparan sulfate (HS) in endothelial basement membrane, as well as better characterized proteases. Heparanase, an endo-.beta.-D-glucuronidase initially detected in B16 melanoma cells, has been described as a Mr 96,000 glycoprotein with pl of 5.2, and has been immunolocalized to the cell surface and cytoplasm. We have utilized a polyacrylamide-gel-based HS degrdn. assay to demonstrate that KNRK, a rat kidney fibroblast cell line transformed by v-K-ras, exhibits HS-degrading activity similar to that of B16F10 mouse melanoma cells. To immunoselect heparanase-expressing clones from a KNRK-cell-specific .lambda.gt11 cDNA library, we have also prepd. a rabbit anti-serum directed against a putative amino-terminal peptide of B16F10 cellular heparanase. Lysogens from one clone expressed a .beta.-galactosidase fusion protein whose staining with peptide anti-serum was inhibited by competition with excess peptide. Dideoxy-mediated sequencing of the insert termini of this recombinant revealed that it represents a rat homolog of Mr 94,000 glucose-regulated protein (GRP94/endoplasmin), a mol. chaperone that contains the exact amino-terminal sequence previously attributed to heparanase. Our results call into question the specificity of this peptide sequence, as well as previous immunolocalization studies of heparanase carried out using such anti-sera.

L8 ANSWER 5 OF 7 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1992:190176 CAPLUS
DOCUMENT NUMBER: 116:190176
TITLE: **Antibodies**, kits, and methods for

immunochemical localization of **heparanase** in mouse and human melanomas, and characterization of melanoma heparanase
INVENTOR(S): Nicolson, Garth L.; Nakajima, Motowo; Jin, Li
PATENT ASSIGNEE(S): University of Texas System, USA
SOURCE: PCT Int. Appl., 82 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9119197	A1	19911212	WO 1991-US3832	19910530 <--
W: AU, CA, JP				
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, NL, SE				
AU 9182317	A1	19911231	AU 1991-82317	19910530 <--
AU 641269	B2	19930916		
EP 532695	A1	19930324	EP 1991-913555	19910530 <--
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE				
JP 05509403	T2	19931222	JP 1991-512410	19910530 <--

PRIORITY APPLN. INFO.: US 1990-530869 19900531
WO 1991-US3832 19910530

AB **Antibodies** to a glycosaminoglycan endoglycosidase (esp. **heparanase**), as well as kits and methods employing the **antibodies**, are disclosed. **Antibodies** against an N-terminal **heparanase** peptide are produced. These antibodies are used for the detection of heparan sulfate endoglycosidase in human and murine tumors. Purifn. of melanoma heparanase is described. A hemocyanin-coupled **heparanase**-derived peptide was used as an immunogen for **antibody** prodn. Also described is prepn. and reactivity of various substrates (e.g. desulfated or desulfated and acetylated heparan sulfate) with melanoma heparanase. The anti-**heparanase antibodies** of the invention stained metastatic melanoma cells, but did not stain surrounding tissue.

L8 ANSWER 6 OF 7 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 5
ACCESSION NUMBER: 1990:529983 CAPLUS
DOCUMENT NUMBER: 113:129983
TITLE: Immunochemical localization of heparanase in mouse and human melanomas
AUTHOR(S): Jin, Li; Nakajima, Motowo; Nicolson, Garth L.
CORPORATE SOURCE: M. D. Anderson Cancer Cent., Univ. Texas, Houston, TX, 77030, USA
SOURCE: International Journal of Cancer (**1990**), 45(6), 1088-95
CODEN: IJCNAW; ISSN: 0020-7136
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Heparanase, an endo-.beta.-D-glucuronidase, has been assocd. with melanoma metastasis. Polyclonal **antibodies** directed against the murine N-terminal **heparanase** peptide detected a Mr .apprx.97,000 protein on SDS-PAGE of mouse melanoma and human melanoma cell lysates. In an indirect immunocytochem. study, human A375-SM and mouse B16-BL6 melanoma cells were stained with the anti-**heparanase antibodies**. Heparanase antigen was localized in the cytoplasm of permeabilized melanoma cells as well as at the cell surface of unpermeabilized cells. Immunohistochem. staining of frozen sections from syngeneic mouse lungs contg. micrometastases of B16-BL6 melanoma demonstrated heparanase localized in metastatic melanoma cells. Similar studies using frozen sections of malignant melanomas resected from patients indicated that heparanase is localized in invading melanoma cells. These studies suggest that (a) the N-terminus of the heparanase

mol. in mouse and human is antigenically related; (b) heparanase antigens are localized at the cell surface and in the cytoplasm of metastatic human and mouse melanoma cells; and (c) heparanase antigens are enriched in invasive and metastatic murine and human melanomas in vivo.

L8 ANSWER 7 OF 7 CAPLUS COPYRIGHT 2002 ACS DUPLICATE 6
ACCESSION NUMBER: 1987:117981 CAPLUS
DOCUMENT NUMBER: 106:117981
TITLE: Soluble antigen induces T lymphocytes to secrete an endoglycosidase that degrades the heparan sulfate moiety of subendothelial extracellular matrix
AUTHOR(S): Fridman, Rafael; Lider, Ofer; Naparstek, Yaakov; Fuks, Zvi; Vlodavsky, Israel; Cohen, Irun R.
CORPORATE SOURCE: Dep. Radiat., Hadassah Univ. Hosp., Jerusalem, 91120, Israel
SOURCE: Journal of Cellular Physiology (**1987**), 130(1), 85-92
CODEN: JCLLAX; ISSN: 0021-9541
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The antigen-mediated induction of heparanase, an endoglycosidase capable of degrading heparan sulfate from the subendothelial extracellular matrix (ECM), was investigated in a rat T lymphocyte cell line reactive against the basic protein (BP) of myelin. It was found that nonactivated T lymphocytes could be induced to express heparanase activity following exposure to sol. but not to ECM-bound BP. The induction of heparanase was immunolog. specific and independent of the presence of syngeneic or allogeneic antigen-presenting cells (APC). However, anti-IA **antibodies** inhibited **heparanase** expression. Sol. BP induced secretion of heparanase into the culture medium within minutes, despite inhibition of protein synthesis. Cell lysates of T lymphocytes contained heparanase activity. Thus, T lymphocytes secrete a preformed heparanase following exposure to specific antigen.

=> dis his

(FILE 'HOME' ENTERED AT 13:53:38 ON 21 NOV 2002)

FILE 'MEDLINE, CAPLUS, EMBASE, BIOSIS' ENTERED AT 13:54:35 ON 21 NOV 2002
L1 406 S PECKER I?/AU OR VLODASKY I?/AU OR FRIEDMAN Y?/AU OR PERETS T?
L2 60 S L1 AND HEPARANASE
L3 12 S L1 AND (HEPARANASE (10N) ANTIBOD?)
L4 9 DUP REM L3 (3 DUPLICATES REMOVED)
L5 79 S (HEPARANASE (10N) ANTIBOD?)
L6 70 S L5 NOT L4
L7 18 S L6 AND PD<19970902
L8 7 DUP REM L7 (11 DUPLICATES REMOVED)